THE UNIVERSITY
OF MICHIGAN
JUN 10 1960
ENGINEERING
LIBRARY

АВТОМАТИКА
И
ТЕЛЕМЕХАНИКА

Volume 20, No. 10

October 1959

SOVIET INSTRUMENTATION AND
CONTROL TRANSLATION SERIES

Automation and Remote Control

(The Soviet Journal *Avtomatika i Telemekhanika* in English Translation)

■ This translation of a Soviet journal on automatic control is published as a service to American science and industry. It is sponsored by the Instrument Society of America under a grant in aid from the National Science Foundation, continuing a program initiated by the Massachusetts Institute of Technology.

SOVIET INSTRUMENTATION AND CONTROL TRANSLATION SERIES

Instrument Society of America Executive Board

John Johnston, Jr.
President
Henry C. Frost
Past President
Dr. Ralph H. Tripp
President-Elect-Secretary
Thomas C. Wherry
Dept. Vice President
Glen G. Gallagher
Dept. Vice President
Adelbert Carpenter
Dept. Vice President
Nathan Cohn
Dept. Vice President
John C. Koch
Treasurer
Nelson Gildersleeve
Dist. I Vice President
Charles A. Kohr
Dist. II Vice President
John R. Mahoney
Dist. III Vice President
George L. Kellner
Dist. IV Vice President
Milton M. McMillen
Dist. V Vice President
Glenn F. Brockett
Dist. VI Vice President
Thomas H. Pierson
Dist. VII Vice President
John A. See
Dist. VIII Vice President
Robert C. Mann
Dist. IX Vice President
Joseph R. Rogers
Dist. X Vice President
John J. McDonald
Dist. XI Vice President

Headquarters Office

William H. Kushnick
Executive Director
Charles W. Covey
Editor, ISA Journal
George A. Hall, Jr.
Assistant Editor, ISA Journal
Herbert S. Kindler
Director, Tech. & Educ. Services
Ralph M. Stotsenburg
Director, Promotional Services
William F. Minnick, Jr.
Director, Public Relations

ISA Publications Committee

Nathan Cohn, *Chairman*
Jere E. Brophy, Enoch J. Durbin, George R. Feeley, Richard W. Jones, George A. Larsen, Thomas G. MacAnespie, John E. Read, Joshua Stern, Frank S. Swaney, Richard A. Terry

Translations Advisory Board of the Publications Committee

Jere E. Brophy, *Chairman*
T. J. Higgins, S. G. Eskin, G. Werbizky

■ This translation of the Soviet Journal *Avtomatika i Telemekhanika* is published and distributed at nominal subscription rates under a grant in aid to the Instrument Society of America from the National Science Foundation. This translated journal, and others in the Series (see back cover), will enable American scientists and engineers to be informed of work in the fields of instrumentation, measurement techniques, and automatic control reported in the Soviet Union.

The original Russian articles are translated by competent technical personnel. The translations are on a cover-to-cover basis, permitting readers to appraise for themselves the scope, status,and importance of the Soviet work.

Publication of *Avtomatika i Telemekhanika* in English translation started under the present auspices in April, 1958, with Russian Vol. 18, No. 1 of January, 1957. Translations of Vols. 18 and 19 have now been completed. Translation of Vol. 20, No. 1 will be published in January, 1960, and the twelve issues of Vol. 20 will be published in English translation by October, 1960.

Transliteration of the names of Russian authors follows the system known as the British Standard. This system has recently achieved wide adoption in the United Kingdom, and is being adopted in 1959 by a large number of scientific journals in the United States.

All views expressed in the translated material are intended to be those of the original authors, and not those of the translators, nor the Instrument Society of America.

Readers are invited to submit communications on the quality of the translations and the content of the articles to ISA headquarters. Pertinent correspondence will be published in the "Letters" section of the ISA Journal. Space will also be made available in the ISA Journal for such replies as may be received from Russian authors to comments or questions by American readers.

1959 Volume 20 Subscription Prices:

Per year (12 issues), starting with Vol. 20, No. 1

General: United States and Canada	$35.00
Elsewhere	38.00
Libraries of nonprofit academic institutions:	
United States and Canada	$17.50
Elsewhere	20.50
Single issues to everyone, each	$ 6.00

1957 Volume 18 and 1958 Volume 19 issues also available. Prices upon request.

See back cover for combined subscription to entire Series.

Subscriptions and requests for information on back issues should be addressed to the:

Instrument Society of America
313 Sixth Avenue, Pittsburgh 22, Penna.

Translated and printed by Consultants Bureau Enterprises, Inc.

Volume XX No. 10 – October 1959

English Translation Published June 1960

Automation and Remote Control

The Soviet Journal Avtomatika i Telemekhanika in English Translation

Reported circulation of the Russian original 8,000.

Avtomatika i Telemekhanika is a Publication of the Academy of Sciences of the USSR

EDITORIAL BOARD
as Listed in the Original Soviet Journal

Corr. Mem. Acad. Sci. USSR V. A. Trapeznikov, *Editor in Chief*
Dr. Phys. Math. Sci. A. M. Letov, *Assoc. Editor*
Academician M. P. Kostenko
Academician V. S. Kulebakin
Corr. Mem. Acad. Sci. USSR B. N. Petrov
Dr. Tech. Sci. M. A. Aizerman
Dr. Tech. Sci. V. A. Il'in
Dr. Tech. Sci. V. V. Solodovnikov
Dr. Tech. Sci. B. S. Sotskov
Dr. Tech. Sci. Ya. Z. Tsypkin
Dr. Tech. Sci. N. N. Shumilovskii
Cand. Tech. Sci. V. V. Karibskii
Cand. Tech. Sci. G. M. Ulanov, *Corresp. Secretary*
Eng. S. P. Krasivskii
Eng. L. A. Charikhov

See following page for Table of Contents.

Copyright by Instrument Society of America 1960

AUTOMATION AND REMOTE CONTROL

Volume 20, Number 10 October 1959

CONTENTS

	PAGE	RUSS. PAGE
On Evaluating the Quality of Nonlinear Automatic Systems with Random Interference. E. P. Popov	1281	1313
L. S. Pontryagin Maximum Principle in the Theory of Optimum Systems. I. L. I. Rozonoér	1288	1320
The Problem of Optimizing Systems which Contain Essentially Nonlinear Elements. E. P. Merkulova	1303	1335
Optimal Control of an Object with Two Controlling Stimuli. E. A. Rozenman	1314	1345
A Method of Synthesizing Linear Sampled-Data Automatic Control Systems to Accord with Dynamic Criteria. L. N. Volgin	1319	1350
AC Servo Systems in which the Signal Depends on the Error and its Derivative. N. P. Vlasov	1329	1357
Graphicoanalytic Method of Computing Steady-State Speed of Pneumatic Drives. V. N. Dmitriev and V. I. Chernyshev	1335	1366
On One Method of Designing Two-Input Electronic Function Generators. V. B. Smolov	1342	1374
On the Determination of Certain Errors of Analog Computers. N. L. Sosenskii	1349	1381
Choosing Flexible Feedback Parameters for Magnetic Amplifiers by Means of Integral Estimates. L. V. Safris	1359	1392
On Increasing Telemetering Effectiveness. N. V. Pozin	1369	1403
Switching Circuit Operation During Transition Periods. V. N. Roginskii	1374	1409
On One Method of Analyzing Complex Switching Circuits. V. F. D'yachenko	1382	1417
Experimental Determination of Power Positioning Servo System Characteristics. M. B. Tumarkin	1391	1426
The Use of Magnetic Amplifiers in Integrating Circuits. S. B. Negnevitskii	1396	1431
Chronicle		
Scientific Seminar on Pneumohydraulic Automata. A. I. Semikova	1401	1435

ON EVALUATING THE QUALITY OF NONLINEAR AUTOMATIC SYSTEMS WITH RANDOM INTERFERENCE

E. P. Popov

(Leningrad)

The method of statistical linearization [1, 2] and its further development from the aspect of harmonic linearization for self-oscillating systems [3, 4] are the most effective means for an approximate investigation of the dynamics of automatic systems with typical nonlinearities (relay, saturation, insensitivity zone, cubic dependence, etc). Some other methods are described in [5]. However, the solution of the given problem, as was shown there, is still in an initial stage of development.

A new application of a combination of the methods of statistical and harmonic linearization to the problem of investigating the influence of random interference on the stability and dynamic properties of nonlinear automatic control systems is given in this article. It is known, for example, that a decrease of the amplification factor of the useful signal under the influence of noise is observed in the presence of saturation nonlinearity (see [6], page 278). It is evident that if such an amplifier is included in a closed automatic system, the latter can substantially change its dynamic properties in relation to the useful signal under the influence of random interference up to the point of possible loss of stability, ultimately. In connection with this, an attempt to develop an approximate theory of the above-mentioned phenomena is undertaken here, and is illustrated with an example of the calculation of a concrete automatic system.

1. Initial Equations

Let a nonlinear automatic control system, having an arbitrary structure, be described by the differential equation

$$Q(p)\,x + R(p)\,F(x) = S_1(p)\,f_1(t) + S_2(p)\,f_2(t), \tag{1.1}$$

where F (x) is the given nonlinearity, f_2 (t) is random interference with zero mathematical expectation, f_1 (t) is a regular external perturbation, which slowly changes in comparison with the interference, and Q (p), R (p), S_1 (p) and S_2 (p) are operators with constant coefficients (p = d/dt), whose form depends on the different distributions of the variable $\underline{x}$ and the external perturbations f_1 (t) and f_2 (t).

Let us seek a solution in the form

$$x = x_M(t) + x_R(t), \tag{1.2}$$

where x_m (t) is the regular component or mathematical expectation (useful signal), x_R (t) is the central random component (interference at the input to the nonlinearity).

Considering the interference f_2 (t) as a stationary random process, given by its spectral density $s_f(\omega)$, we will also consider the solution for the component x_R (t) as a stationary random process. This supposition is based on the slow change of the mathematical expectation x_m (t), in so far as it can be considered constant for possible "periods" of change of x_R (t) corresponding to the given frequency spectrum $s_f(\omega)$; that is, the case is examined

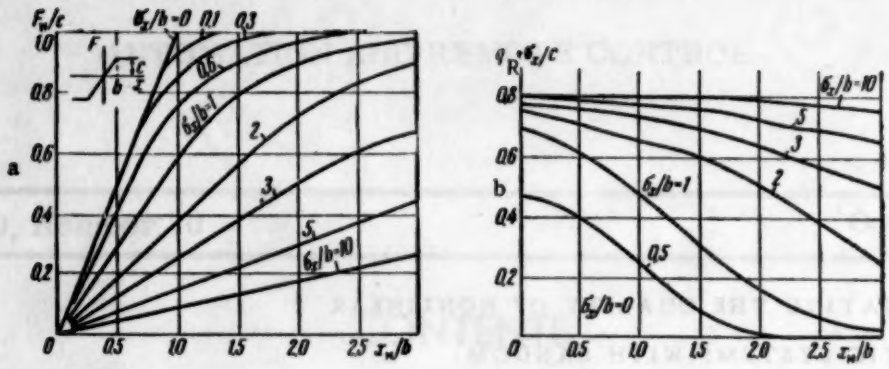

Fig. 1.

when the interference frequencies are high in comparison to the rate at which the useful signal can be changed.

By using, as in the case of harmonic linearization, the property of a filter in the linear part of the system, we will approximately characterize the random processes in the nonlinear system (1.1) by the two first statistical moments: the mathematical expectation (mean value) and the dispersion (or mean square deviation). In correspondance with this, we will produce the following statistical linearization of the nonlinearity:

$$F(x) = F_M + q_R x_R, \tag{1.3}$$

where F_m is the regular component (mathematical expectation) of the nonlinear function, q_R is the amplification factor of the random component in the nonlinear link, where for single-valued nonlinearities

$$F_M = M[F(x_M + x_R)] = \int_{-\infty}^{+\infty} F(x_M + x_R)\, w(x)\, dx$$

for the normal distribution law

$$w(x) = \frac{1}{\sigma_x \sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x - x_M}{\sigma_x}\right)^2},$$

where σ_x is the mean square value of the random component x_R. We will select the simpler second method [1] for determination of q_R:

$$q_R = \frac{1}{\sigma_x^2} \int_{-\infty}^{+\infty} F(x_M + x_R)\, x w(x)\, dx.$$

Analogous formulas can also be written for loop nonlinearities.

As a result, we have ready expressions for any typical nonlinearity

$$F_M(x_M, \sigma_x), \tag{1.4}$$
$$q_R(x_M, \sigma_x), \tag{1.5}$$

which are depicted graphically, as is shown, for example, in Fig. 1, for saturation nonlinearity [1].

Putting (1.2) and (1.3) in the given equation of the nonlinear system (1.1), we will separate it into two equations:

$$Q(p)\, x_M + R(p)\, F_M = S_1(p)\, f_1(t), \tag{1.6}$$
$$Q(p)\, x_R + R(p)\, q_R\, x_R = S_2(p)\, f_2(t), \tag{1.7}$$

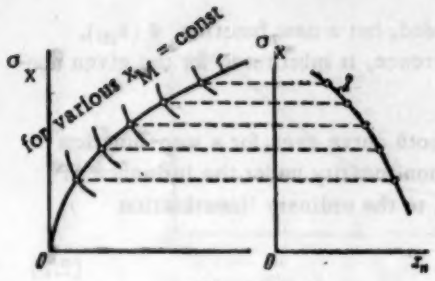

Fig. 2.

respectively for the regular component (mathematical expectation), which is the useful signal, and for the random component (central), which corresponds to interference.

The solutions of Eqs. (1.6) and (1.7) are associated with one another in a nonlinear manner, according to (1.4) and (1.5); this is essential for the ultimate manifestation of the nonlinear properties of the given system.

On the basis of the equation for the random component (1.7), and knowing the spectral density of the external interference $s_f(\omega)$, it is possible to determine the dispersion of the interference at the input of the nonlinear link σ_x^2 according to the known formula

$$\sigma_x^2 = \frac{1}{2\pi}\int_{-\infty}^{+\infty}\left|\frac{S(j\omega)}{Q(j\omega)+R(j\omega)\,q_R(x_M,\,\sigma_x)}\right|^2 s_f(\omega)\,d\omega$$

or

$$\sigma_x^2 = aI_n(x_M,\,\sigma_x), \tag{1.8}$$

where a is a constant factor (including the ordinarily given quantity σ_f) placed outside the integral sign, and

$$I_n = \frac{1}{2\pi}\int_{-\infty}^{+\infty}\frac{G(\omega)}{|H(j\omega)|^2}\,d\omega, \tag{1.9}$$

in which $G(\omega)$ includes the square of the modulus of $S(j\omega)$ and the part of the numerator of $s_f(\omega)$, which depends on ω, while $|H(j\omega)|^2$ is the whole denominator of the integrand; the denominator of $s_f(\omega)$, transformed in an appropriate manner, is also included in the term $|H(j\omega)|^2$. In the literature (see for example, [2], page 643) there are prepared expressions for I_n up to n = 7, where n is the degree of the polynomial $H(j\omega)$.

From Eq. (1.8), it is possible to find the mean square value of the interference σ_x at the input of the nonlinear link, as a function of the magnitude of the useful signal

$$\sigma_x(x_M) \tag{1.10}$$

graphically, in a manner similar to the solution for A (x^*) in [7]. In order to do this, we write

$$\sigma_x^2 = \zeta, \quad aI_n(x_M,\,\sigma_x) = \zeta$$

instead of (1.8), and we construct a parabola 1 (Fig. 2) on the plane ζ, σ_x from the first equation, and a series of curve segments $\sigma_x(\zeta)$ for various x_m = const from the second equation. Their intersections with curve 1, when translated to the plane x_m, σ_x (curve 2 in Fig. 2) give the solution sought for (1.10).

2. Investigation of the Properties of a Nonlinear System in Relation to the Useful Signal in the Presence of Random Interference

Having determined the dependence $\sigma_x(x_m)$, in such a manner we will place it in expression (1.4); as a result, σ_x is eliminated, and we get the function

$$F_M = \Phi(x_M), \tag{2.1}$$

which can serve as a characteristic for determining the passage of the useful signal through the nonlinear link in the presence of random interference. Having placed it in (1.6), we obtain the nonlinear equation for the useful signal of the system as a whole in the form

$$Q(p)\,x_M + R(p)\,\Phi(x_M) = S_1(p)\,f_1(t), \tag{2.2}$$

which differs from the true Eq. (1.1) in that here the interference is discarded, but a new function $\Phi(x_m)$, which was found for the useful signal and accounts for the effect of interference, is substituted for the given nonlinearity F (x).

It appears that the function $\Phi(x_m)$ ordinarily has the shape of a smooth curve even for a step-function nonlinearity F (x) (relay, loop, etc) – the effect is that of smoothing the nonlinearity under the influence of random interference. Therefore it is possible to subject it, in many cases, to the ordinary linearization

$$\Phi(x_M) = k_H x_M, \tag{2.3}$$

where k_H is determined either graphically as the slope of the curve $\Phi(x_m)$, or analytically:

$$k_H = \left(\frac{d\Phi}{dx_M}\right)_{x_M=0}. \tag{2.4}$$

Then the nonlinear equation for the useful signal of the system (2.2) is replaced by the linear equation:

$$[Q(p) + k_H R(p)]\, x_M = S_1(p) f_1(t), \tag{2.5}$$

which significantly simplifies all further investigation of the dynamic properties of the system in relation to the useful signal (in the presence of random interference), since any known methods of linear automatic control theory can be used to this end.

However, in spite of the use of linear methods, all the essential nonlinear properties of the system are manifested, due to the particular method, presented above, of determining the amplification factor of the useful signal in the nonlinear link k_H in the presence of random interference. Indeed, the quantity k_H, like the function $\Phi(x_m)$, is not simply a given number distinct from the amplification factors of ordinary linear links. According to (1.4) and (1.8), the quantities k_H and $\Phi(x_m)$ depend, firstly, on the spectral density of the external random interference $s_f(\omega)$, on the basis of its mean square level σ_f and frequency spectrum; secondly, on the shape of the nonlinearity F (x); and thirdly, on the structure and on all other parameters of the system itself.

All the statistical and dynamic properties of the system, which were determined from Eq. (2.5), including the system stability, will depend on these factors. Their dependence on the level of the external interference is an especially important distinction from ordinary linear systems (see example below).

The dependence of the quantity k_H on the structure and on all the parameters of the system (manifested through the magnitude of σ_x, from which k_H is determined) is very important to consider in synthesizing the system, that is, in the choice of the structure and parameters of the system, according to Eq. (2.5), for desirable dynamic properties. If, for example, the magnitude of any parameter, k, of the system is selected, then it is necessary preliminarily to determine the dependence $k_H(k)$, which is not difficult to do in the method for finding k_H described above (see example below).

The method of determining the quantity k_H can be significantly simplified in the particular case when the linear part of the system, determined by the transfer function

$$W_L(p) = \frac{R(p)}{Q(p)}, \tag{2.6}$$

does not pass the frequency spectrum corresponding to the given spectral density of the external interference $s_f(\omega)$, that is, it is considered in the particular given case that $R(j\omega)/Q(j\omega) \approx 0$ is in the limits of the spectrum of the random component $x_R(t)$. Then, dividing Eq. (1.7) by Q (p) and considering the above-mentioned property, we obtain a simple formula for the dispersion of the interference at the input of the nonlinear link:

$$\sigma_x^2 = \frac{1}{2\pi}\int_{-\infty}^{+\infty}\left|\frac{S(j\omega)}{Q(j\omega)}\right|^2 s_f(\omega)\, d\omega, \tag{2.7}$$

which, unlike (1.8), does not depend on the magnitude of the useful signal x_m.

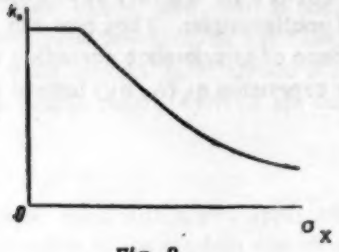

Fig. 3.

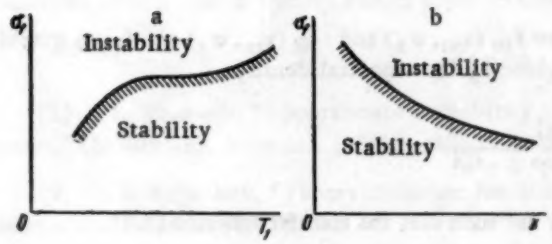

Fig. 4.

According to (2.1) and (1.4), the function $\Phi(x_m)$ will concur with $F_m(x_m, \sigma_x)$ in this case, that is, it will be determined directly from a graph of the type shown in Fig. 1a, at an appropriate value of σ_x (2.7). Consequently, the amplification factor of the useful signal in the nonlinear link k_H also will be determined simply as the tangent of the slope angle of the appropriate curve of $F_m(x_m)$ in Fig. 1a, for a given σ_x or

$$k_{\text{н}} = \left(\frac{\partial F_{\text{м}}}{\partial x_{\text{м}}}\right)_{x_{\text{м}}=0} \tag{2.8}$$

directly from expression (1.4).

In other systems it is often observed that although the linear part (2.6) does pass the frequency spectrum $s_f(\omega)$, it is not passed by some other separate block (most often, the object being controlled). Then a simpler nonlinear equation of the system without this block is formed for the interference, instead of (1.7). In such a case, one is obliged to use the general method of determining σ_x according to (1.8), but for a simpler equation (lower order).

In both of the particular cases described, which are often encountered, the simplification of the solution is concerned only with the determination of how $\Phi(x_m)$ or k_H depends on the interference, after which the complete equation of the system (2.2) or (2.5) is used in all cases for an investigation of the properties of the system in relation to the useful signal.

If the linearization of the function $\Phi(x_m)$ in the form (2.3) is undesirable for some reason and it is necessary to evaluate the effect of the virtual curvature of the function $\Phi(x_m)$, one is obliged to investigate the nonlinear equation (2.2) directly with any of the known nonlinear methods of control theory. In particular, it is possible to find the stability of the system from Eq. (2.2) by the method of harmonic linearization or harmonic balance [8, 9], and also to find the self- and forced-oscillations in relation to the useful signal [10, 11], and to evaluate the transients [12, 13]. In order to do this, it is essential to produce a harmonic linearization of the function $\Phi(x_m)$ for $x_m = A \sin \psi$ in the form

$$\Phi(x_{\text{м}}) = q_{\text{h}}(A, \sigma_x)\, x_{\text{м}},$$

where

$$q_{\text{h}} = \frac{1}{\pi A}\int_0^{2\pi} F_{\text{м}}(A\sin\phi,\ \sigma_x)\sin\phi\, d\phi. \tag{2.9}$$

Here $\Phi(x_m)$ is replaced by the expression $F_m(x_m, \sigma_x)$, since it is necessary that the linear part (2.6) does not pass the higher harmonics of the useful signal, in order to use harmonic linearization; this means that the whole interference spectrum must not be passed. * Therefore, the quantity σ_x in expression (2.9) is determined from the simple formula (2.7).

As a result, we get the following harmonically linearized equation for investigating the properties of the system in relation to the useful signal, instead of (2.2):

$$[Q(p) + q_{\text{h}}(A, \sigma_x)\, R(p)]\, x_{\text{м}} = S_1(p)\, f_1(t) \tag{2.10}$$

at the particular values $f_1 = 0$ and $f_1 = B \sin \Omega t$.

*In those cases when the interference frequencies are higher than the possible self-oscillation frequencies in relation to the useful signal.

Ready expressions for the coefficient of harmonic linearization $q_h(A, \sigma_x)$, obtained after statistical linearization which was done earlier, were given in [3] for different types of nonlinearities. They now can be used for the calculation of a system in relation to the useful signal in the presence of interference according to Eq. (2.10) with exactly the same methods as in [8-13], but with the use of a new expression $q_h(A, \sigma_x)$ instead of $q(A)$; σ_x is determined from formula (2.7).

3. Example

A nonlinear automatic system is described by the equation

$$p^3(T_2p+1)x + (k_2k_{OS}p^2 + T_1kk_1p + kk_1)F(x) = kp^3(T_1p+1)f_2(t)$$

with a saturation nonlinearity F(x), for which the expressions $F_m(x_m, \sigma_x)$ and $q_R(x_m, \sigma_x)$ are shown graphically in Fig. 1. The random external interference $f_2(t)$ is given by the spectral density

$$s_f(\omega) = \frac{2\beta\sigma_f^2}{(\omega_1^2 - \alpha^2\omega^2)^2 + \mu^2\omega^2}.$$

The parameters of the system and the spectral density are such that the transfer function (2.6)

$$W_L(p) = \frac{k_2k_{OS}p^2 + T_2kk_1p + kk_1}{p^3(T_2p+1)}$$

does not pass the frequency spectrum $s_f(\omega)$. It is necessary to determine the stability of the system in relation to the parameters $\underline{k}$ and T_1, depending on the interference level σ_f.

We find geometrically from Fig. 1a the quantity k_H in the function from σ_x (Fig. 3). According to (2.5), we write the characteristic equation of the system in relation to the useful signal

$$p^3(T_2p+1) + (k_2k_{OS}p^2 + T_1kk_1p + kk_1)k_n(\sigma_x) = 0,$$

the condition of the system stability according to the Hurwitz criterion will be

$$k_n(\sigma_x) > \frac{1}{T_1(k_2k_{OS} - T_2kk_1T_1)}. \tag{3.1}$$

We have from formula (2.7)

$$\sigma_x^2 = \frac{1}{2\pi}\int_{-\infty}^{+\infty}\left|\frac{k(T_1j\omega+1)}{T_2j\omega+1}\right|^2 \frac{2\beta\sigma_f^2}{(\omega_1^2 - \alpha^2\omega^2)^2 + \mu^2\omega^2}\,d\omega,$$

whence

$$\sigma_x = k\sigma_f\sqrt{2\beta I_3}, \tag{3.2}$$

where

$$I_3 = \frac{1}{2\pi}\int_{-\infty}^{+\infty}\frac{(T_1^2\omega^2+1)\,d\omega}{|\alpha^2T_2(j\omega)^3 + (\alpha^2+\mu T_2)(j\omega)^2 + (\omega_1^2T_2+\mu)j\omega + \omega_1^2|^2} =$$

$$= \frac{\alpha^2 + \mu T_2 - T_1^2\omega_1^2}{2\mu\omega_1^2(\alpha^2 + \mu T_2 + T_2^2\omega_1^2)}.$$

Using the stability condition (3.1), the expression for σ_x (3.2), and the dependence $k_H(\sigma_x)$ (Fig. 3), we can plot graphs of the system stability in relation to any desired parameters, depending on the interference level. In particular, the stability limits have the form shown in Fig. 4a and b, in relation to the parameters T_1 and $\underline{k}$. Here the essential dependence of the limits of system stability (in relation to the useful signal) on the random

interference level can be seen; this is not observed in linear systems. The possibility of loss of stability at a definite interference level, or, in any case, the essential dependence of all statistical and dynamic properties of the system on the external interference is an important peculiarity, which must be taken into account when nonlinear automatic control systems are designed.

SUMMARY

The paper deals with a new application of a combination of statistic and harmonic linearizing methods to the investigation of the random disturbance effect on stability and dynamic properties of nonlinear automatic control systems. In this connection there was made an attempt to develop the rough theory of these phenomena with a control system design given to illustrate it.

LITERATURE CITED

[1] I. E. Kazakov, "Approximate probability analysis of the operating precision of essentially nonlinear systems," (In Russian), Avtomat. i Telemekh. Vol. 17, No. 5 (1956). *

[2] V. S. Pugachev, "Theory of random functions and its application to automatic control problems," (In Russian), Gostekhizdat (1957).

[3] A. A. Pervozvanskii, "Approximate method of investigating self-oscillating systems in the presence of random perturbations," [In Russian], Izv. Akad. Nauk SSSR, Otd. Tekhn. Nauk, No. 3 (1958).

[4] A. A. Pervozvanskii, "Self-oscillating systems in the presence of slowly changing external perturbations," [In Russian], Izv. Akad. Nauk SSSR, Mekhan. i mashinostr. No. 1 (1959).

[5] V. V. Solodovinkov, "Analysis and synthesis of nonlinear automatic control systems in the presence of random perturbations," [In Russian], in the set Itogi Nauki, 1st edition, chap. 5 (1957).

[6] V. A. Besekerskii, V. P. Orlov, L. V. Polonskaya and S. M. Fedorov, "Designing low power follower systems," Sudpromgiz (1958).

[7] E. P. Popov, "Effect of vibration interference on the stability and dynamic properties of nonlinear automatic control systems," [In Russian], Izv. Akad. Nauk SSSR, Otd. tekhn. Nauk, energet. i avtom., No. 4 (1959).

[8] K. Magnus, VDI-Forschungsh, No. 451 (1955).

[9] E. P. Popov, "On the discrimination of the stability regions of nonlinear automatic systems on the basis of harmonic linearization," [In Russian], Izv. Akad. Nauk SSSR, Otd. tekhn. nauk, energet. i avtom. No. 1 (1959).

[10] L. S. Gol'dfarb, "On some nonlinearities in control systems," [In Russian], Avtomat. i Telemekh. No. 5 (1947).

[11] E. P. Popov, "Approximate investigation of self- and forced-oscillations in high order nonlinear systems," [In Russian], Izv. Akad. Nauk SSSR, Otd. tekhn. nauk, No. 5 (1954).

[12] E. P. Popov, "Approximate investigation of transients in nonlinear automatic systems by the method of harmonic linearization," [In Russian], Izv. Akad. Nauk SSSR, Otd. tekhn. nauk, No. 9 (1956).

[13] A. A. Voronov, "Approximate determination of the establishment of self-oscillations in certain nonlinear automatic control systems," [In Russian], Avtomat. i Telemekh. Vol. 18, No. 7 (1957). *

Received May 6, 1959

*See English translation.

L. S. PONTRYAGIN MAXIMUM PRINCIPLE IN THE THEORY OF OPTIMUM SYSTEMS. I

L. I. Rozonoér
(Moscow)

Questions are discussed, which are associated with the proof and use of the L. S. Pontryagin maximum principle in the theory of optimum systems. The work also contains some new results. The problem of optimization for the case of a free right end of a trajectory is examined in the first part of the work. In the second part, the maximum principle is formulated for boundary conditions of a more general type. The connection between the method of dynamic programing and the maximum principle is established in the third part; a method of solving optimization problems in linear discrete systems is given, and a number of considerations concerning the use of the maximum principle in the solution of a definite class of problems associated with the theory of dynamic accuracy of control systems are also discussed.

The problem of creating systems, which are optimum in some prescribed sense, is one of the most important problems of automatic control quality theory. Questions about obtaining optimum laws of control were first seriously posed, apparently, in the theory of the motion of a rocket (see, for example, [17-19]). However, the results obtained here are standard from the viewpoint of general control theory.

The majority of the most important and interesting practical problems lead to nonclassical variation problems, which require new mathematical methods for their solution. The circumstance comes from the fact that the variables in the equation of a real system are (or should be) limited in some way or other. The first most general and important results in this sphere concern theories of systems which insure minimum time of control processes. In 1952-1955, the bases of the theory of processes in linear systems, optimum in relation to rapid action, were developed, mainly in the works of A. A. Feld'baum and A. Ya. Lerner. In particular, the solution of the problem of synthesizing systems, optimum in relation to time, for the case of a linear aperiodic object of arbitrary order with one controlling perturbation, was given in [1-2]. A detailed review of results obtained up to 1956 can be found in [3-5].

In 1956, a principle, leading to the solution of the general problem of finding a control process, optimum for rapid action, was hypothesized by L. S. Pontryagin on the basis of the results of work performed by him, V. G. Boltyanskii, and R. V. Gamkrelidze [6-7]. This principle, which received the name "Maximum Principle," was verified at first for individual types of systems, and in particular, was proved in [8] for the case of linear systems. V. G. Boltyanskii [9] fully proved that the Pontryagin maximum principle was a necessary condition for optimality in relation to rapid action. R. V. Gamkrelidze [8-10] proved theorems of existence and uniqueness and examined the problem of synthesizing controls for linear systems, optimum for rapid action. In [11], the maximum principle was extended to the general case of minimizing an arbitrary functional of the integral function of variable systems. In [12], there is a detailed presentation of basic results obtained by L. S. Pontryagin and his associates.

In examining questions concerned with the theory of optimum systems, it is necessary to note the numerous

works of R. Bellman, which are systematically presented in [13] (see also [14]). The method of "Dynamic Programming," developed by R. Bellman, gives a new tool for the solution of nonclassical variational problems, which are closely associated with the Pontryagin maximum principle.

The works of N. N. Krasovskii [15, 16] are also concerned with general questions on the theory of optimum systems; discrete systems, described by finite difference equations, are examined in [16].

The short review presented above is concerned only with papers connected with the maximum principle in some way or other. Hence, many interesting papers were not mentioned.

This paper contains a presentation of the most essential problems in the theory of automatic control which are associated with the proof and use of the Pontryagin maximum principle. The presentation of the majority of the known results (in particular, the proof of the maximum principle) differs somewhat from that in the original works. References to the original works are not given in the text as a rule, so as not to complicate the presentation. Some new results are also contained in this work; a formula for the increment of a functional is based on these results; and also, the sufficient conditions of optimality, formulation of boundary conditions for a number of problems, the connection between the maximum principle and the method of dynamic programming, the formulation and proof of a principle for optimum processes in linear discrete systems, analogous to the maximum principle, result from this formula. The work is published in the form of three separate articles. The optimization problem for the case of a free right end of a trajectory is examined in the first part. In the second part, the maximum principle is formulated for the general case of arbitrary boundary conditions, and the problem of synthesizing optimum systems is discussed. In the third part, the connection between the maximum principle and the method of dynamic programming is established; a method of solving some problems of optimization in discrete systems is given; also, a number of considerations concerning the use of the maximum principle in the solution of a definite class of problems connected with the theory of dynamic precision of control systems are presented (in particular, B. V. Bulgakov's problems about the accumulation of deviations for nonlinear systems).

The most complicated proofs have been placed in appropriate appendices.

1. Statement of the Problem

Let us examine an object of automatic control, which has r controlling elements and is described by a system of differential equations of the n-th order:

$$\dot{x}_i = f_i(x_1, \ldots, x_n; u_1, \ldots, u_r; t) \qquad (i = 1, \ldots, n), \tag{1}$$

where $x_1, \ldots, x_n$ are the parameters of the object, and $u_1, \ldots, u_r$ are the positions of the controlling elements. At each moment, the positions of the controlling elements $u_1(t), \ldots u_r(t)$ must satisfy the inequalities

$$\varphi_j(u_1(t), \ldots, u_r(t)) \leqslant 0 \qquad (j = 1, \ldots, m), \tag{2}$$

which reflect the restrictions which are imposed on the control system. In the future, we will use a geometrical terminology, and we will consider an r-fold space R $(u_1, \ldots, u_r)$. Some point (vector) u of the space R corresponds to any given combination of the positions of the controlling elements $(u_1, \ldots, u_r)$. The vector $u(t) = (u_1(t), \ldots, u_r(t))$, given as a time function, will be called the "control" of the object. The control u (t) will be considered piecewise continuous, that is, each of the functions $u_k(t)$ can have a finite number of discontinuities of the first type at the end of an interval of time.

The inequalities (2) isolate some closed* set of points, whose boundaries are given by the equalities $\varphi_j(u_1, \ldots, u_r) = 0$, in the space R. This set will be designed by U, and the inequalitities (2) can be briefly written in the form

$$u(t) \in U, \tag{2'}$$

that is, the vector u (t) must belong to the closed set U of the space R at any given moment. In this case, our

*The set is called closed if it includes its own boundary. The set being examined is closed in so far as conditions (2) permit a sign of the equalities such that the point u can be found on the boundary of the set.

choice of the possible controls is limited by condition (2') and by the requirement of piecewise continuity of the functions $u_1(t),\ldots, u_r(t)$. The control u (t), which obeys these two conditions, will be called admissible. When speaking about controls from now on, we will have admissible controls in mind.

We will examine an n-fold phase space X $(x_1,\ldots, x_n)$ of the parameters of the object, along with the space R $(u_1,\ldots, u_r)$, in which the control can be changed. The state of the object is described then by the point (vector) x $(x_1,\ldots, x_n)$ of the space X. If the control u (t) and the initial state of the system $x^0 = (x_1^0,\ldots, x_n^0)$ are given, then the behavior of the system (trajectory x (t) in the phase space X) is uniquely determined from Eqs. (1).

The basic problem of the theory of optimum systems is the choice of a control u (t) such that the behavior of the system will be "best" in some prescribed sense. We will introduce a number of examples of precise statements of problems, in which the sense of the concept "optimum control" is made concrete each time.

It is possible to require that the control u (t) be so selected that the system will go from one fixed state x^0 to another fixed state x^1 in minimum time. The control u (t), which satisfies this requirement, is called optimum for rapid action. The time of transition of the system from state x^0 to state x^1, which determines the criteria of optimality in the given problem, is a functional* given on the controls u (t). The requirement of optimality of the control u (t) means minimization of the value of the functional.

Let us examine another problem. Let it be required that a control u (t) be chosen such that in a given time T the system will go from an initial state x^0 into a state in which one of the variables (for example, x_1) will become as large as possible, and the remaining coordinates will be fixed values. This problem arises, for example, in the calculation of the control law of a rocket, or in putting an artificial satellite of the Earth into a given orbit [20], if it is required that the horizontal velocity of the rocket be maximum at a fixed height and at a given time, for zero vertical velocity and for a given flight range. In the present example, the functional, which determines the criteria of optimality, is the value of the generalized coordinate $x_1(T)$ (horizontal velocity of the rocket) at the time t = T. In a number of cases, there is no necessity to fix the final values of all the remaining coordinates $x_2,\ldots, x_n$. If, for example, only the height and not the flight range plays a role, there is no necessity to impose any restrictions on the latter. Then an analogous problem about the maximum of the functional $x_1(T)$ arises, but it is not necessary for all the remaining coordinates to assume prescribed values. Let us examine, finally, one more problem. It is necessary to choose a control u (t) in such a manner that the integral

$$S = \int_0^T F(x_1, \ldots, x_n;\ u_1, \ldots, u_r)\, dt$$

becomes minimum (or maximum) during the time of movement T. Here again, different variations are possible. It can be required that the initial and final states of the system be fixed beforehand (for example, in the problem about the flight of an airplane from one fixed point to another, with the least fuel consumption; the function F (x, u) represents fuel consumed per unit time, and the integral represents the full consumption of fuel during the flight). In other cases, the final state of the system does not have any value; it is only important that the integral become the maximum (or minimum) value of all those possible (for example, if it is required that the aggregate being controlled give the maximum quantity of production in the absence of any other conditions; in this case, the function F (x, u) represents instantaneous producitivity, and the integral represents the full production output). The time T, during which optimization of the system is required, can also be either fixed beforehand, or not fixed. If the airplane moves between two points on a schedule, then its flight time is fixed, and minimization of fuel consumption must be required for this time, fixed beforehand. If time does not matter, but only fuel economy is important, then the time must not be fixed beforehand.

The number of examples of such optimization problems can be increased further. However, all these superficially different problems prove to be closely associated mathematically, and require completely analogous methods of investigation. Moreover, one of them can quickly to transformed into others.

*A functional is a correspondence established between functions (or combinations of functions) and numbers. In the given example, a number — the transition time from x^0 to x^1 — corresponds to each control u (t) [that is, the combination of functions $u_1(t),\ldots, u_r(t)$] that insures that the system gets to x^1.

We will show, for example, that the problem of optimizing a system with respect to an integral leads to the problem of optimizing with respect to coordinates. Actually, let it be necessary to optimize the magnitude

$\int_0^T F(x_1, \ldots, x_n;\ u_1, \ldots, u_r;\ t)\,dt$ (the time T, and also the final values of the coordinates can be either fixed, or free) in an object described by the differential equations (1), with restrictions on the control (2). We will introduce a new variable

$$x_{n+1}(t) = \int_0^t F(x_1, \ldots, x_n;\ u_1, \ldots, u_r;\ t)\,dt,\ x_{n+1}(0) \equiv x_{n+1}^0 = 0.$$

Then one more differential relation can be added to system (1)

$$\dot{x}_{n+1} = F(x_1, \ldots, x_n;\ u_1, \ldots, u_r;\ t),$$

and the problem of optimizing the integral leads to the problem of optimizing the n + 1 st coordinate $x_{n+1}(T)$ at the final moment of time.

The problem about the optimization of some function from the final value of the coordinates $\Phi[x_1(T), \ldots, x_n(T)]$ also leads easily to the problem about optimizing the final value of one coordinate, if the function Φ is differentiable.

To show this, it is sufficient to introduce a new coordinate,

$$x_{n+1}(t) = \Phi[x_1(t), \ldots, x_n(t)],\quad x_{n+1}(0) \equiv x_{n+1}^0 = \Phi(x_1^0, \ldots, x_n^0)$$

and to add the relation to system (1)

$$\dot{x}_{n+1} = \sum_{s=1}^{n} \frac{\partial \Phi(x_1, \ldots, x_n)}{\partial x_s} f_s(x_1, \ldots, x_n;\ u_1, \ldots, u_r;\ t).$$

Let us further note that the problem of optimization with "integral" restrictions of the type

$$\int_0^T F(x_1, \ldots, x_n;\ u_1, \ldots, u_r;\ t)\,dt \leqslant A$$

easily leads to ordinary statements of the problem. Introducing the variable $x_{n+1}(t)$ and adding the relation

$$x_{n+1} = F(x_1, \ldots, x_n;\ u_1, \ldots, u_r;\ t),$$

to (1), we arrive at the ordinary problem, with the additional restriction on the final value of the n + 1st coordinate: $x_{n+1}(T) \leq A$.

The problem about the minimum transition time of the system from one fixed state to another is also a particular case of the problem about the minimization of one coordinate.

Indeed, by combining equation

$$\frac{dx_{n+1}}{dt} = 1,$$

with (1), we may convince ourselves that minimization of time means minimization of the coordinate x_{n+1}. In this case, the final values of the remaining coordinates are given, and the time T, evidently, should not be fixed.

Hence, a broad class of optimization problems leads to the problem concerning optimization of one

coordinate of the system. It will be more convenient for us, however, to examine a somewhat more general problem (although one which easily leads to the problem about optimization of one coordinate). To wit, we will pose the problem of optimizing a linear function of the final values of all the coordinates of the system, that is, the quantities $S = \sum_{k=1}^{n} c_k x_k(T)$, where c_k are certain constants. The quantities S can be treated as the scalar product of the vector $x(T) = [x_1(T), \ldots, x_n(T)]$, which gives the end point of the trajectory x (t), and the given vector $c = (c_1, \ldots, c_k)$. In other words, we require optimization of the projection of the vector x (T) on the given direction $(c_1, \ldots, c_k)$. By the same token, the requirement of maximum (or minimum) size of S means that we attempt to move the system as much "farther" in the direction of the vector $\underline{c}$ as possible (or as much "farther" in the direction of the vector $-c$ as possible).

As we have seen, various restrictions can be imposed on the final state of the system (that is, on the position of the final point of the trajectory x (T) in the phase space X). In general, these restrictions can be formulated as a requirement of a transfer of the system to some fixed set G of the phase space X, described by some combination of equalities and inequalities. In particular, if the final values of some coordinates of the system are precisely fixed, and the remainder free, then the set G is a linear manifold in the space X ("linear", "plane," etc).

Hence, we will examine the following problem. It is necessary to select a control u (t) from the group of admissible controls which transfer the system (1) from the point $x(T_0) \equiv x^0$ to the fixed closed set G of the phase space, in such a way that the sum $S = \sum_{i=1}^{n} c_i x_i(T)$ at the given time t = T will assume a minimum (or maximum) value.

In order to be definite, we have formulated a problem in which the control time $T-T_0$ is fixed. However, we will see consequently that in those problems where the time is not given beforehand (in particular, in the problem about minimizing the transition time), completely analogous results obtain.

The control u (t) which makes the functional $S = \sum_{1}^{n} c_i x_i(T)$, minimum (maximum) will be called min-optimum (max-optimum) in S, for brevity.

Mathematically, the problem examined by us is a generalization of Mayer's classical problem in the calculus of variations. However, the presence of restrictions on the positions of the controlling elements, leading to the requirement that the controls u (t) must belong to the closed set U, generally speaking does not permit one to use the methods of the classical calculus of variations. Therefore, the solution of the problem formulated above requires the development of new mathematical methods.

It should further be noted that the theory of optimum systems is not to any degree exhausted by the solution of the problem formulated above. Firstly, various complications in the statement of the problem are possible. In particular, accounting for restrictions, which can be imposed on the phase coordinates $x_1, \ldots, x_n$, has great practical value. Secondly, in a number of cases problems arise which do not lead to optimization of the final value of the coordinates of the system and which require the solution of variational problems for functionals of another type (for example, the problem about the minimum of the maximum deviation of some coordinate from a given level). These and analogous questions are not discussed in this article.

2. The Problem with the Free Right End of the Trajectory

In this section we will examine a particular case of our basic problem. To wit, we will consider here that the right end of the trajectory x (T) is "free", that is, no restrictions are imposed on the final values of the coordinates (the set G occupies all of the space X). The results, which pertain to the case being examined, prove to be completely analogous to the results obtained in the solution of the general problem, differing from the latter only in details which concern the boundary conditions (transversality conditions) and at the same time permit more simple proofs. Moreover, problems of this type are important in themselves, for example, in a

number of questions about the theory of the dynamic precision of control systems or in solving problems of control in mathematical economics. Therefore it is well to examine this simpler problem separately.

Below, we will formulate and discuss results leading to the solution of the problem being examined. Full proofs of the theorems are in the appendix.

We will bring into consideration a combination of n time functions $p_1(t),\ldots p_n(t)$, which form a variable vector $p(t) = (p_1(t),\ldots, p_n(t))$, which has a direction, at time $t = T$, opposite to the direction of the vector $c = (c_1,\ldots, c_n)$, that is

$$p_i(T) = -c_i \quad (i = 1, \ldots, n). \tag{3}$$

We assume here, for simplicity, that the modului of the vectors p (T) and c are equal. The variables $p_i(t)$ are subject to a set of differential equations of the following form

$$\dot{p}_i(t) = -\sum_1^n p_s \frac{\partial f_s(x_1, \ldots, x_n; u_1, \ldots, u_r; t)}{\partial x_i} \qquad (i = 1, \ldots, n), \tag{4}$$

by the same token, we give a law which the vector p (t) can be changed in accordance with. Let us note that if any control u (t) is given, the vector p (t) is uniquely determined from Eqs. (4), where conditions (3) play the role of boundary conditions. Indeed, if the control u (t) is known, then, by solving the system of equations

$$\dot{x}_i = f_i(x, u, t) \qquad (i = 1, \ldots, n),$$

for which the initial conditions $[x_i(T_0) = x_i^0,\ i = 1,\ldots, n)]$ are given, we can find the functions $x_i(t)$. Putting these functions in (4), we would have a set of linear differential equations with variable coefficients and with boundary conditions (3), for the magnitudes of $p_i(t)$. The solution of these equations would permit us to determine p (t) as a function of time.

The vector p (t) will be called an "impulse" which maps points at time t.

At each time t, let us examine the scalar product $\sum_1^n p_i(t)\dot{x}_i(t)$ of the vector p (t) and the velocity vector $\dot{x}$ (t) which maps the points. Bringing Eqs. (1) to our attention, we will express this quantity in the form of a function of $2n + r + 1$ variables $x_1,\ldots, x_n;\ p_1,\ldots, p_n;\ u_1,\ldots, u_r;\ t$:

$$H(x, p, u, t) \equiv \sum_{s=1}^n p_s f_s(x_1, \ldots, x_n; u_1, \ldots, u_r; t). \tag{5}$$

By using the function H (x, p, u, t), which plays a basic role in all further discussion, Eqs. (1) and (4) can be written in the form

$$\dot{x}_i = \frac{\partial H}{\partial p_i}, \quad \dot{p}_i = -\frac{\partial H}{\partial x_i} \qquad (i = 1, \ldots, n) \tag{6}$$

with the boundary conditions

$$x_i(T_0) = x_i^0, \quad p_i(T) = -c_i \qquad (i = 1, \ldots, n). \tag{7}$$

Let us turn our attention to the fact that Eqs. (6) concur in form with the canonical equations of Hamilton in analytical mechanics. The function H is analogous to the Hamiltonian, and the vector p (t) to the impulse vector, which explains the name given by us to the vector p (t). The connection between the equations of the theory of optimum systems and the equations of mechanics is discussed also in appendix 4.

Now let u (t) be some admissible control, and $x^u(t)$, $p^u(t)$ be the position and impulse which maps the points at time t for the control u (t). Let us put the values of the variables $x_i^u(t)$, $p_i^u(t)$ in the function

H (x, p, u, t). Then we will obtain some function

$$K(t, u_1, \ldots, u_r) \equiv H(x^u(t), p^u(t), u, t),$$

which is a function of r variables $u = (u_1, \ldots, u_r)$ for every fixed time t, that is, points which belong to the closed set U of the space R. We will say that the control u (t) satisfies the maximum (minimum) condition if the function K (t, u) reaches an absolute maximum (minimum) in the set U at any time t ($T_0 \leq t \leq T$), for values of the variables equal to the values of the control at the same time, that is, at $u_k = u_k(t)$ ($k = 1, \ldots, r$). It appears that fulfilling the maximum (minimum) condition for u (t) is a necessary condition for min-optimality (max-optimality) of the control u (t); this leads to the theorem expressing the Pontryagin maximum principle for our problem. *

Theorem 1. If the control u (t) is min-optimum (max-optimum) in $S = \sum_1^n c_i x_i(T)$, then it satisfies the maximum (minimum) condition.

Hence, the optimum control u (t) must be chosen at each moment so that the function H (x^u (t), p^u (t), u, t), which is a scalar product of the vector p^u (t) and the velocity vector $\dot{x}^u$ (t), is minimized (maximized). Therefore, the geometrical sense of the theorem is that the optimum control tries to "drive away" the image point optimally in some direction determined by the vector p^u (t), at each moment. At the time t = T, when $p^u(T) = -c$, the direction of "maximum dispersal" concurs with the direction of the vector c if we are seeking the maximum of the functional $S = \sum_1^n c_i x_i(T)$, and in the opposite direction to c if the minimum of the functional is sought. This property of the optimum control becomes especially clear in the following example, where the equations of the object have the form:

$$\dot{x}_i = f_i(u_1, \ldots, u_r) \quad (i = 1, \ldots, n),$$

that is, the right halves of the equations do not depend on the coordinates. Forming the differential equations for the impulse p^u (t), we find that the vector p^u (t) is constant in our case, and is equal to

$$p^u(t) = -c \qquad (T_0 \leqslant t \leqslant T).$$

at any time.

Indeed, $\partial f_s / \partial x_i \equiv 0$, and it follows from Eqs. (4) [or (6)] that $\dot{p}^u(t) = 0$ and, in view of conditions (3) [or (7)], we obtain the equality which is written. Therefore, the direction of "maximum dispersal" is always constant in the given case, and at any time, the control is chosen so that the maximum possible velocity is imparted to the direction of the vector c for max-optimality, and the maximum possible velocity in the direction of the vector − c for min-optimality. Naturally, such a simple control law does not prevail in the general case, since if the right halves of the equations of the object depend on the coordinates, then interval "forces" will act on the image point, and these must be taken into account. The "internal" forces are accounted for by introducing a variable vector p^u (t).

In the future, we will omit the index u in the designations p^u (t), x^u (t).

The maximum principle permits one to form a set of differential equations whose solution determines the

---

*We note that fulfillment of the maximum condition is essential if the functional $S = \sum_1^n c_i x_i(T)$ is to be minimum (and vice-versa). This is connected with the fact that the final value of the vector p (t) was defined as a vector opposite to the vector $c = (c_1, \ldots, c_n)$. One can postulate that p (T) = + c, but in this case the formulation of the theorems would be the reverse of analogous formulations assumed in [6-12], which would lead to confusion.

optimum trajectory and optimum direction. Indeed, the quantities $u_1, \ldots, u_r$ are determined from the maximum (minimum) condition, at any time, as functions of the variables $x_1, \ldots, x_n$; $p_1, \ldots, p_n$; t:

$$u_k = \psi_k(x_1, \ldots, x_n;\ p_1, \ldots, p_n;\ t) \qquad (k = 1, \ldots, r).$$

Putting the expressions obtained in the right halves of Eqs. (6), we will get a set of 2n differential equations with 2n unknown $x_1(t), \ldots, x_n(t)$; $p_1(t), \ldots, p_n(t)$, whose solution, under the boundary conditions (7), will give the optimum trajectory x (t) and impulse p (t). Knowing x (t) and p (t), we obtain the optimum control u (t).

Let us take the simplest example. Let it be required to find the control u (t), which makes the integral

$\frac{1}{2}\int_0^T (x^3 + u^2)\,dt$, maximum, if the equation of the object has the form:

$$\dot{x} = -ax + u, \quad x(0) = x^0,$$

and no restrictions are imposed on the control u (t) (that is, the set U occupies all of the space R— in the given case, the whole real axis $-\infty < u < +\infty$).

Introducing the variables $x_1(t) \equiv x(t)$, $x_2(t) \equiv \frac{1}{2}\int_0^t (x_1^2 + u^2)\,dt$, we get the system

$$\dot{x}_1 = -ax_1 + u, \quad \dot{x}_2 = \frac{1}{2}x_1^2 + \frac{1}{2}u^2.$$

The functional which must be minimized is $S = x_2(T)$, that is, $c_1 = 0$, $c_2 = 1$. According to equality (5), we write the function

$$H(x,\ p,\ u) \equiv -ap_1x_1 + \frac{1}{2}p_2x_1^2 + p_1u + \frac{1}{2}p_2u^2$$

and the equations for the impulse (second group of Eqs. (6))

$$\dot{p}_1 = ap_1 - p_2x_1, \quad \dot{p}_2 = 0$$

with the kind of boundary conditions of equalities (7)

$$x_1(0) = x^0, \quad x_2(0) = 0; \quad p_1(T) = 0; \quad p_2(T) = -1.$$

We notice at once that $p_2(t)$ is constant and $p_2 = -1$ always, so that the function H has the form:

$$H = -ap_1x_1 - \frac{1}{2}x_1^2 + p_1u - \frac{1}{2}u^2.$$

According to the maximum principle, the function H must be maximum in $\underline{u}$ for any values of $\underline{x}$ and $\underline{p}$. In the given case, the maximum of H is reached at the value $u = p_1(t)$. Putting the value of $\underline{u}$ obtained in the equation of the object, we obtain the following equations for the determination of $x_1(t)$ and $p_1(t)$ (the function $x_2(t)$ does not interest us):

$$\dot{x}_1 = -ax_1 + p_1, \quad \dot{p}_1 = x_1 + ap_1$$

with boundary conditions $x_1(0) = x^0$, $p_1(T) = 0$. The equations are easily integrated:

$$x_1(t) = C_1e^{\lambda t} + C_2e^{-\lambda t}, \quad p_1(t) = D_1e^{\lambda t} + D_2e^{-\lambda t},$$

where $\lambda = \sqrt{a^2 + 1}$ is the root of the characteristic equation, and the constants C_1, C_2, D_1, D_2 depend on x^0. Since $u(t) = p_1(t)$ the optimum control is also determined.

The problem which was examined can evidently be easily solved without the use of the maximum principle, since there are no restrictions on the controls. The given example was introduced only to illustrate the procedure of the solution coming from the maximum principle.

In a very broad class of practically important problems it appears that the maximum of the function $K(t, u) \equiv H[x(t), p(t), u, t]$ is reached at the boundary of the set U, and the optimum control takes its boundary values. In particular, this circumstance always occurs if the functions $u_1(t), \ldots, u_r(t)$ enter the system of equations of the problem linearly. For example, let the equations of the problem have the form:

$$\dot{x}_1 = f_1(x_1, x_2) + u_1, \qquad \dot{x}_2 = f_2(x_1, x_2) + u_2,$$

and the functions $u_1(t)$ and $u_2(t)$ be limited in modulus so that: $|u_1| \leq M_1$, $|u_2| \leq M_2$ (the set U is a rectangle). Then, forming the function

$$H(x, p, u) \equiv p_1 f_1(x_1, x_2) + p_2 f_2(x_1, x_2) + p_1 u_1 + p_2 u_2,$$

we find that the maximum of H is reached at $u_1 = M_1 \operatorname{sign} p_1$, $u_2 = M_2 \operatorname{sign} p_2$. Indeed, H is maximum in u_1 and u_2 when u_1 and p_1, u_2 and p_2 respectively have identical signs (so that the products p_1u_1 and p_2u_2 are positive); in this case, u_1 and u_2 reach a maximum in modulus.* Therefore, the functions $u_1(t)$ and $u_2(t)$ can only "switch" from one boundary position to another, no matter how $p_1(t)$ and $p_2(t)$ change with time; $p_1(t)$ and $p_2(t)$ are determined only at the moment of switching.

Theorem 1, formulated above, which expresses the Pontryagin maximum principle, gives only the necessary conditions of optimality. It is easy to show examples in which the control satisfies the maximum (minimum) condition, but never the less is not optimum. From the viewpoint of the remaining conditions, it appears that the control, which is subject to the maximum (minimum) conditions is optimum, generally speaking, only in a definite local sense, if some additional conditions are not fulfilled, which are connected with the "sharpness" of the maximum (minimum) of the function H [21]. However, for a system whose controls are linear relative to the coordinates

$$\dot{x}_i = \sum_1^n a_{is}(t)\, x_s + \varphi_i(u_1, \ldots, u_r), \quad i = 1, \ldots, n. \tag{8}$$

the following theorem obtains.

Theorem 2. The necessary and sufficient condition of min-optimality (max-optimality) of the control u (t) in system (8) is the fulfillment of the maximum (minimum) condition for u (t).

It must be noted, that if the maximum (minimum) condition is not sufficient, generally speaking, even so, it permits one to determine uniquely the optimum control in the majority of practically important cases. Indeed, if the optimum control exists (which is often evident from physical considerations), and the control satisfying the maximum (minimum) condition is unique, then this unique control is optimum.

APPENDIX

1. Increment of the functional during a change of the control. Let x_i and p_i satisfy the equations

$$\dot{x}_i = \frac{\partial H}{\partial p_i}, \quad \dot{p}_i = -\frac{\partial H}{\partial x_i} \quad (i = 1, \ldots, n), \tag{9}$$

*We designate the function

$$z = \begin{cases} 1 & \text{for } y > 0, \\ -1 & \text{for } y < 0, \end{cases}$$

by the symbol $z = \operatorname{sign} y$, which is not defined when $y = 0$. If at any time $p_1(t) = 0$ and $p_2(t) = 0$, then H does not, in fact, depend on u_1 and u_2, and the maximum principle does not permit one to find the value of the optimum control (see appendix 3.3).

where

$$H(x,\ p,\ u,\ t) \equiv \sum_1^n p_i f_i(x, u, t). \tag{10}$$

The functions f_i ($i = 1,\ldots, n$) will be considered continuous in all of the arguments (x, u, t), and to have continuous partial derivatives in the arguments (x, u), up to and including the second derivative.

The equality

$$I(x,\ p,\ u) \equiv \int_{T_*}^{T} \left[\sum_1^n p_i \dot{x}_i - H(x, p, u, t)\right] dt = 0, \tag{11}$$

follows from (9) and (10), and is justified if the first group of Eqs. (9) is satisfied under arbitrary controls and boundary conditions. We will choose some control u (t) $\in$ U and examine its increment δu (t). Let x (t), and p (t) be a solution of (9) for the control u (t) and some boundary conditions, and x (t) + δx (t), p (t) + δp (t) be a solution of (9) for the control u (t) + δu (t) with the same boundary conditions. Let us examine the difference $\Delta \equiv I(x + \delta x, p + \delta p, u + \delta u) - I(x, p, u)$, in which $\Delta = 0$ always. It follows from (11) that

$$\Delta = \int_{T_*}^{T} \sum_1^n (p_i \delta \dot{x}_i + \delta p_i \dot{x}_i)\, dt + \int_{T_*}^{T} \sum_1^n \delta p_i \delta \dot{x}_i dt -$$

$$- \int_{T_*}^{T} [H(x + \delta x,\ p + \delta p,\ u + \delta u,\ t) - H(x,\ p,\ u,\ t)]\, dt. \tag{12}$$

We note that the functions x (t), p (t), δx (t), δp (t) are continuous, and their derivatives exist and are piecewise continuous.

We will integrate by parts:

$$\int_{T_*}^{T} p_i \delta \dot{x}_i dt = p_i \delta x_i \Big|_{T_*}^{T} - \int_{T_*}^{T} \dot{p}_i \delta x_i dt; \qquad \int_{T_*}^{T} \delta p_i \delta \dot{x}_i dt = \delta p_i \delta x_i \Big|_{T_*}^{T} - \int_{T_*}^{T} \delta x_i \delta \dot{p}_i dt.$$

Hence, keeping (9) in mind, and introducing the vector $y = (y_1, \ldots, y_{2n})$ ($x_i = y_i$, $p_i = y_{n+i}$, $i = 1, \ldots, n$) for brevity, we obtain

$$\int_{T_*}^{T} \sum_1^n (p_i \delta \dot{x}_i + \delta p_i \dot{x}_i)\, dt = \sum_1^n p_i \delta x_i \Big|_{T_*}^{T} + \int_{T_*}^{T} \sum_1^{2n} \frac{\partial H(y, u, t)}{\partial y_s} \delta y_s dt, \tag{13}$$

$$\int_{T_*}^{T} \sum_1^n \delta p_i \delta \dot{x}_i dt = \sum_1^n \frac{1}{2} \delta p_i \delta x_i \Big|_{T_*}^{T} + \frac{1}{2} \int_{T_*}^{T} \sum_1^{2n} \delta \frac{\partial H(y, u, t)}{\partial y_s} \delta y_s dt. \tag{14}$$

Using Taylor's formula, we put the last integral in (12) into the form

$$\int_{T_*}^{T} [H(y + \delta y,\ u + \delta u,\ t) - H(y, u, t)]\, dt = \int_{T_*}^{T} [H(y, u + \delta u, t) - H(y, u, t)]\, dt +$$

$$+ \int_{T_*}^{T} \sum_1^{2n} \frac{\partial H(y, u + \delta u, t)}{\partial y_s} \delta y_s dt + \frac{1}{2} \int_{T_*}^{T} \sum_{s,q=1}^{2n} \frac{\partial^2 H(y + \vartheta_1 \delta y,\ u + \delta u, t)}{\partial y_s \partial y_q} \delta y_s \delta y_q dt, \tag{15}$$

where $0 < \vartheta_1(t) < 1$. Since

$$\frac{1}{2}\delta\frac{\partial H}{\partial y_s}\delta y_s \equiv \frac{1}{2}\frac{\partial H(y+\delta y, u+\delta u, t)}{\partial y_s}\delta y_s - \frac{1}{2}\frac{\partial H(y, u, t)}{\partial y_s}\delta y_s,$$

by putting (13)-(15) in (12), we get

$$\Delta \equiv \sum_1^n \left(p_i + \frac{1}{2}\delta p_i\right)\delta x_i\Big|_{T_0}^{T} - \int_{T_0}^{T}[H(y, u+\delta u, t) - H(y, u, t)]\,dt -$$

$$-\frac{1}{2}\int_{T_0}^{T}\sum_1^{2n}\left[\frac{\partial H(y, u+\delta u, t)}{\partial y_s} - \frac{\partial H(y, u, t)}{\partial y_s}\right]\delta y_s dt -$$

$$-\frac{1}{2}\int_{T_0}^{T}\sum_{s,q=1}^{2n}\frac{\partial^2 H(y+\vartheta_1\delta y, u+\delta u, t)}{\partial y_s \partial y_q}\delta y\,\delta y_q dt +$$

$$+\frac{1}{2}\int_{T_0}^{T}\sum_1^{2n}\left[\frac{\partial H(y+\delta y, u+\delta u, t)}{\partial y_s} - \frac{\partial H(y, u+\delta u, t)}{\partial y_s}\right]\delta y_s dt.$$

Using Taylor's formula once more for the integrand of the last integral, and keeping in mind that $\Delta = 0$, we will have

$$\sum_1^n \left(p_i + \frac{1}{2}\delta p_i\right)\delta x_i\Big|_{T_0}^{T} = \int_{T_0}^{T}[H(y, u+\delta u, t) - H(y, u, t)]\,dt + \eta, \tag{16}$$

where $\eta = \eta_1 + \eta_2$ and

$$\eta_1 = \frac{1}{2}\int_{T_0}^{T}\sum_1^{2n}\left[\frac{\partial H(y, u+\delta u, t)}{\partial y_s} - \frac{\partial H(y, u, t)}{\partial y_s}\right]\delta y_s dt,$$

$$\eta_2 = \frac{1}{2}\int_{T_0}^{T}\sum_{s,q=1}^{2n}\left[\frac{\partial^2 H(y+\vartheta_1\delta y, u+\delta u, t)}{\partial y_s \partial y_q} - \frac{\partial^2 H(y+\vartheta_2\delta y, u+\delta u, t)}{\partial y_s \partial y_q}\right]\delta y_s \delta y_q dt, \tag{17}$$

in which $0 < \vartheta_1 < 1$, $0 < \vartheta_2 < 1$. Fixing the boundary conditions in the form

$$x_i(T_0) = x_i^0, \quad p_i(T) = -c_i \qquad (i = 1, \ldots, n), \tag{18}$$

in view of the fact that $\delta x_i(T_0) = \delta p_i(T) = 0$, we get, finally,

$$\delta S \equiv \sum_1^n c_i \delta x_i(T) = -\int_{T_0}^{T}[H(x, p, u+\delta u, t) - H(x, p, u, t)]\,dt - \eta. \tag{19}$$

2. Evaluation of the remaining term in the formula for the increment of the functional. First, we will get an evaluation of $\delta x_i(t)$ and $\delta p_i(t)$.

It follows from the first group of Eqs. (9) that

$$\delta\dot{x}_i = f_i(x+\delta x, u+\delta u, t) - f_i(x, u, t) \quad (i = 1, \ldots, n).$$

Since the initial conditions are fixed and $\delta x_i(T_0) = 0$, then

$$\delta x_i(t) = \int_{T_0}^{t}[f_i(x+\delta x, u+\delta u, t) - f_i(x, u, t)]\,dt.$$

By using the Lipschitz conditions, which always hold for our assumptions, we obtain the inequalities

$$|\delta x_i(t)| \leqslant \int_{T_0}^{t} K \sum_{1}^{n} |\delta x_s(z)|\, dz + \int_{T_0}^{t} K \sum_{1}^{r} |\delta u_k(z)|\, dz \quad (i = 1, \ldots, n),$$

where K is the maximum of the Lipschitz constants for the functions f_i.

Summing and strengthening the inequalities, we have

$$X(t) \leqslant Kn \int_{T_0}^{t} \sum_{1}^{r} |\delta u_k(z)|\, dz + \int_{T_0}^{t} KnX(z)\, dz,$$

where $X(t) = \sum_{1}^{n} |\delta x_i(t)|$. It is easy to show that the inequality

$$X(t) \leqslant a + \int_{T_0}^{T} bX(z)\, dz$$

follows from the inequality

$$X(t) \leqslant ae^{b(t-T_0)}$$

(a and b are positive constants).

Therefore, putting $a = Kn \int_{T_0}^{T} \sum_{1}^{n} |\delta u_k(z)|\, dz$, $b = Kn$ and substituting t for T, we get

$$X(t) \leqslant Kne^{Kn(T-T_0)} \int_{T_0}^{T} \sum_{1}^{r} |\delta u_k(z)|\, dz \quad (T_0 \leqslant t \leqslant T).$$

Thence, we have analogous inequalities for $\delta x_i(t)$:

$$\delta x_i(t) \leqslant M_1 \int_{T_0}^{T} \sum_{1}^{r} |\delta u_k(z)|\, dz \quad (i = 1, \ldots, n; \quad T_0 \leqslant t \leqslant T),$$

where $M_1 = Kne^{Kn(T-T_0)}$ is a constant which does not depend on $\delta u(t)$.

Examining the second group of Eqs. (9), we obtain for $p_i(t)$

$$\delta p_i = -\left[\frac{\partial H(x + \delta x, u + \delta u, p + \delta p, t)}{\partial x_i} - \frac{\partial H(x, p, u, t)}{\partial x_i}\right],$$

in which $\delta p_i(T) = 0$. Therefore, by now using the evaluation for $\delta x_i(t)$, we get, in a completely analogous fashion,

$$\delta p_i(t) \leqslant M_2 \int_{T_0}^{T} \sum_{1}^{r} |\delta u_k(z)|\, dz \quad (i = 1, \ldots, n, \; T_0 \leqslant t \leqslant T).$$

Hence, the evaluation

$$\delta y_s(t) \leqslant M_1 \int_{T_0}^{T} \sum_{1}^{r} |\delta u_k(z)|\, dz \quad (s = 1, \ldots, 2n, \; T_0 \leqslant t \leqslant T), \tag{20}$$

occurs, where M is a constant which does not depend on $\delta u(t)$. Applying (12) to the remaining members η_1 and η_2 in (9), using the Lipschitz conditions for the functions $\partial H/\partial y_s$ and having the boundedness of the functions $\partial^2 H/\partial y_s \partial y_q$ in mind, we obtain

$$|\eta| \leqslant A\left(\int_{T_*}^{T}\sum_{1}^{r}|\delta u_k(z)|\,\partial z\right)^{\nu} \tag{21}$$

for $\eta = \eta_1 + \eta_2$, without difficulty.

If the function $\delta u(t)$ is different from zero only in the interval $[t_1, t_2]$ of length $t_2 - t_1 = \tau$, then, using the Cauchy-Bunyakovskii (Schwartz) inequality, formula (21) can be re-written in the form

$$|\eta| \leqslant A\tau\int_{t_1}^{t_2}\left(\sum_{1}^{r}|\delta u_k(t)|\right)^2 dt \quad \text{or} \quad |\eta| \leqslant B\tau\int_{t_1}^{t_2}\sum_{1}^{r}\delta u_k^2(t)\,dt \qquad (B = rA). \tag{22}$$

3. Proof of theorem 1. To be definite, let us examine the case when the minimum of the functional S is sought. For any change $\delta u(t)$ of the control u (t), the following must hold;

$$\delta S \geqslant 0. \tag{0}$$

Let us assume that the maximum condition does not hold at $t = t^*$, that is, a vector $u^* = (u_1^*, \ldots, u_r^*) \in U$ exists such that

$$H[x(t^*), p(t^*), u^*, t^*] > H[x(t^*), p(t^*), u(t^*), t^*]. \tag{:}$$

In view of the piecewise continuity of u (t), and the continuity of the functions x (t), p (t), and H (x, p, u, t) some interval $[t', t'']$, which includes t^*, is always found, on which the functions $F_1(t) \equiv H[x(t), p(t), u^*, t]$ and $F_2(t) \equiv H[x(t), p(t), u(t), t]$ are continuous, and consequently, are uniformly continuous (if $t = t^*$ is a discontinuity point of the control u (t), then, by agreeing that $u(t^*) = u(t^* - 0)$, one can take $t' = t^*$). Therefore, the existance of a number $\alpha > 0$ follows from inequality (:), such that the inequality

$$H[x(t), p(t), u^*, t] - H[x(t), p(t), u(t), t] > \alpha \tag{::}$$

holds for all $t \in [t_1, t_2]$, where $[t_1, t_2]$ are sufficiently small intervals which belong to the interval of continuity $[t', t'']$. Let us now examine the varied control u (t), which is equal to the control $\bar{u}(t)$ everywhere outside the interval $[t_1, t_2]$, and equal to u^* at $[t_1, t_2]$. The control $\bar{u}(t) \in U$ is admissible. Then, using (19), (::), and (22b), we get

$$\delta S = -\int_{t_1}^{t_2}[H(y(t), u^*, t) - H(y(t), u(t), t)]\,dt - \eta <$$

$$< -\left[\int_{t_1}^{t_2}\alpha\,dt - |\eta|\right] \leqslant -\int_{t_1}^{t_2}\left[\alpha - \tau B\sum_{1}^{r}\delta u_k^2(t)\right]dt.$$

Since $\sum_{1}^{r}\delta u_k^2(t)$ is bounded, it is always possible to choose a sufficiently small $\tau = t_2 - t_1$, for which the integrand is positive and $\delta S < 0$, which satisfies condition (0). The case of the maximum of the functional S is examined analogously. The theorem is proved.

4. Proof of theorem 2. The necessity of the maximum (minimum) condition has been established by theorem 1. We will prove the sufficiency of this condition in the case of linear systems. The Hamiltonian of the system has the form:

$$H \equiv \sum_{i,k=1}^{n} a_{ik}(t)\,p_i x_k + \sum_{1}^{n} p_i \varphi_i(u),$$

and the impulse satisfies the relations

$$\dot{p}_i = -\sum_1^n a_{ki}(t)\, p_k \quad (i = 1, \ldots, n).$$

Let us examine the remaining terms (17) in formula (19) and prove that $\eta \equiv 0$. Indeed, $\partial^2 H/\partial x_i \partial x_j \equiv 0$, $\partial^2 H/\partial p_i \partial p_j \equiv 0$, and the quantities $a_{ij} \equiv \partial^2 H/\partial p_i \partial x_j$ clearly do not depend on $\underline{x}$ and $\underline{p}$. Therefore $\eta_2 = 0$. In the expression for η_1, the differences $\partial H(y, u + \delta u, t)/\partial x_i - \partial H(y, u, t)/\partial x_i$ disappear identically, since $\dfrac{\partial H}{\partial x_i} \equiv -\sum_1^n a_{ki}(t)\, p_k$ clearly does not depend on $\underline{u}$. On the other hand, the change of the impulse $\delta p(t)$ satisfies the relations

$$\delta\dot{p}_i = -\sum_1^n a_{ki}(t)\, \delta p_k \quad (i = 1, \ldots, n),$$

which do not depend on δx and δu. But since the boundary conditions for the impulse are fixed and $\delta p(T) = 0$, then $\delta p(t)$ is identically equal to zero for any $\underline{t}$. Therefore, $\eta_1 = 0$. Hence, the formula for the increment of the functional takes the form:

$$\delta S = -\int_{T_0}^{T} [H(y, u + \delta u, t) - H(y, u, t)]\, dt.$$

If the maximum (minimum) condition holds, then the integrand is not positive (not negative), and $\delta S \geq 0$ ($\delta S \leq 0$), Q.E.D.

LITERATURE CITED

[1] A. A. Fel'dbaum, "Optimum processes in automatic control systems," [In Russian], Avtomat. i Telemekh. Vol. 14, No. 5 (1953).

[2] A. A. Fel'dbaum, "On the synthesis of optimum systems with the help of a phase space," [In Russian], Avtomat. i Telemekh. Vol. 16, No. 2 (1955).

[3] A. A. Fel'dbaum, "On the question of synthesizing optimum automatic control systems," [In Russian], Transactions of the Second All-Union Conference on Automatic Control Theory, Vol. 2, Izd-vo. Akad. Nauk SSSR (1955).

[4] A. Ya. Lerner, "Construction of rapid-acting systems with restrictions on the coordinates of the object," [In Russian], see [3].

[5] Itogi nauki. Tekhn. nauki L., "Problems on the theory of nonlinear automatic regulation and control systems" [In Russian], edited by Ya. Z. Tsypkin, Izd-vo. Akad. Nauk SSSR, (chap. 4, paragraph 3, 4) (1957).

[6] V. G. Boltyanskii, R. V. Gamkrelidze and L. S. Pontryagin, "On the theory of optimum processes," [In Russian], Doklady Akad. Nauk SSSR, Vol. 110, No. 1 (1956).

[7] L. S. Pontryagin, "Some mathematical problems arising in connection with the theory of optimum automatic control systems," [In Russian], Session of the Academy of Sciences of the USSR on Scientific Problems of Automating Industry, 15-20 October (1956); "Basic problems of automatic regulation and control," Izd-vo. Akad. Nauk SSSR (1957).

[8] R. V. Gamkrelidze, "On the theory of optimum processes in linear systems" [In Russian], Doklady Akad. Nauk SSSR, Vol. 116, No. 1 (1957).

[9] V. G. Boltyanskii, "The maximum principle in the theory of optimum processes" [In Russian], Doklady Akad. Nauk SSSR, Vol. 119, No. 6 (1958).

[10] R. V. Gamkrelidze, "The theory of processes which are optimum for rapid-action in linear systems" [In Russian], Izv. Akad. Nauk SSSR, ser. matem., Vol. 22, No. 4 (1958).

[11] R. V. Gamkrelidze, "On the general theory of optimum processes," [In Russian], Doklady Akad. Nauk SSSR, Vol. 123, No. 2 (1958).

[12] L. S. Pontryagin, "Optimum control processes" [In Russian], Uspekhi Matem. Nauk, Vol. 13, issue 1 [85] (1959).

[13] R. Bellman, Dynamic Programming. [Princeton University Press, Princeton, N.J., 1957].

[14] Modern Mathematics for Engineers, [Russian translation] edited by E. F. Bakkenbach (chap. 10), [Izd-vo. inostr. lit-ry, M., (1958)].

[15] N. N. Krasovskii, "On the theory of optimum control" [In Russian], Avtomat. i Telemekh. Vol. 18, No. 11 (1957) [See English translation].

[16] N. N. Krasovskii, "On one problem of optimum control" [In Russian], Prikiad. Matem. i Mekhan. Vol. 21, issue 5 (1957).

[17] A. B. Kosmodem'yanskii, "Extremum problems for points of variable mass" [In Russian], Doklady Akad. Nauk SSSR, Vol. 53, No. 1 (1946).

[18] H. S. Tsien and R. C. Evans, "Optimum Thrust Programming for a Sounding Rocket," J. Amer. Rocket Soc. Vol. 21, No. 5 (1951).

[19] A. Hibbs, "Optimum Burning Program for Horizontal Flight," J. Amer. Rocket Soc. Vol. 22, No. 4 (1952).

[20] D. E. Okhotsimskii and T. M. Eneev, "Some variation problems connected with launching artificial Earth satellites" [In Russian], Uspekhi Fiz. Nauk Vol. 62, issue 1 (1957).

[21] L. I. Rozonoér, "On the sufficient conditions of optimality" [In Russian], Doklady Akad. Nauk SSSR, Vol. 127, No. 3 (1959).

Received June 5, 1959

THE PROBLEM OF OPTIMIZING SYSTEMS WHICH CONTAIN ESSENTIALLY NONLINEAR ELEMENTS

E. P. Merkulova

(Moscow)

A method is given for finding the optimal weight function of the linear portion of a stationary system which contains an arbitrary number of essentially nonlinear elements. The problem is solved for various methods of connecting the linear portions and the nonlinear elements. The input is a stationary random function of time consisting of signal plus noise. The linearization of the nonlinear elements is carried out by a statistical method.

The increased accuracy required of automatic control processes in the design of optimal systems makes it ever more compelling to turn our attention to the problem of taking nonlinear characteristics into account.

The literature contains a number of works [1-3] which present various methods for designing optimal nonlinear systems. However, either the final results are given in a form which is difficult to use in practical calculations [1], or the nonlinearity considered is a continuous function [3] or must be given as a power series [2].

In the present work, we present a method of designing optimal dynamic systems which contain essentially nonlinear elements which are either relay-type elements with dead zones or elements with limited zones of linearity.

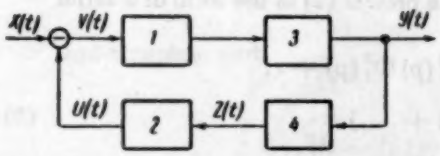

Fig. 1. 1) Nonlinear element No. 1; 2) nonlinear element No. 2; 3) linear portion No. 1 and 4 is linear portion No. 2.

The design of the nonlinear elements is carried out by Kazakov's method [4], i.e., a nonlinear element is replaced by a linear one in such wise that the mathematical expectation and the dispersion of the element's output signal remain as they were previously.

1. Posing of the Problem and the Problem's Solution

We consider a system consisting of any number of linear portions and nonlinear elements which are so connected in a forward circuit and in a feedback path that the system may be represented in the form shown in Fig. 1.

We shall denote by $\Phi(p)$ and $\Psi(p)$ the transfer functions of linear portions No. 1 and No. 2 respectively.

We assume the following with respect to input signal X (t):

1) $X(t) = m(t) + n(t)$;

2) $m(t) = m_0 + m'(t)$ is a stationary random function of time with mathematical expectation m_0, correlation function $R_m(\tau)$ and spectral density $S_m(\omega)$;

3) n (t) is a stationary fandom function of time with zero mathematical expectation, correlation function $R_n(\tau)$ and spectral density $S_n(\omega)$;

4) m (t) and n (t) are noncorrelated;

5) the input signal's spectral density, $S_{\varphi}(\omega) = S_m(\omega) + S_n(\omega)$, is a rational-fractional function of ω.

We shall denote by w (t) the weight function of the system as a whole, and we shall assume that

$$w(t) = 0 \qquad (0 > t > T),$$

where [0, T] is the time interval during which we locate information about the input stimuli, i.e., the system's memory.

After the application of statistical linearization, nonlinear element No. 1 is replaced by coefficients l_0 and l_1, and nonlinear element No. 2 by coefficients k_0 and k_1 (by constant and random signals, respectively). These coefficients are functions of the mathematical expectation and the dispersion of the signals applied to the inputs of the corresponding nonlinear elements.

As a result of these replacements, the system as a whole will have two different transfer functions: one, $\Phi_1(p)$, for the random signals and one, $\Phi_0(p)$, for the constant signals, for which we can write:

$$\Phi_0(p) = \frac{l_0\Phi(p)}{1 + l_0k_0\Phi(p)\Psi(p)}, \quad \Phi_1(p) = \frac{l_1\Phi(p)}{1 + l_1k_1\Phi(p)\Psi(p)}. \tag{1}$$

Since the types of the nonlinear elements are given, the problem reduces to the finding of the optimal transfer function for the linear correcting element, considering the transfer function of the system's invariant portion as being given. We shall assume that $\Psi(p)$ is given. Then, by eliminating $\Phi(p)$ from the two equations in (1), we obtain

$$\Phi_0(p) = \frac{\Phi_1(p)}{q + \nu\Psi(p)\Phi_1(p)}, \tag{2}$$

where

$$q = \frac{l_1}{l_0}, \quad \nu = k_0 - k_1q.$$

By assuming that q is close to unity and ν is close to zero, we can present (2) in the form of a series

$$\Phi_0(p) = \frac{1}{q}\{\Phi_1(p) - \nu\Psi(p)\Phi_1^2(p) + \nu^2\Psi^2(p)\Phi_1^3(p) + \ldots$$
$$\ldots + (-1)^{n-1}\nu^{n-1}\Psi^{n-1}(p)\Phi_1^n(p) + \ldots\}. \tag{3}$$

If we go from transfer functions $\Phi_0(p)$ and $\Phi_1(p)$ to the corresponding weight functions, $w_0(t)$ and $w_1(t)$, we get

$$w_0(t) = \frac{1}{q}\left\{w_1(t) - \nu\int_0^t g_1(t-\theta)\,w_2(\theta)\,d\theta + \ldots\right.$$
$$\left.\ldots + (-1)^{n-1}\nu^{n-1}\int_0^t g_{n-1}(t-\theta)\,w_n(\theta)\,d\theta + \ldots\right\}, \tag{4}$$

where

$$w_n(t_1) = \int_0^{t_1} w_1(t_2)\,dt_2 \int_0^{t_1-t_2} w_1(t_3)\,dt_3 \ldots \int_0^{t_1-t_2-\ldots-t_{n-1}} w_1(t_n)\,w_1(t_1-t_2-\ldots-t_n)\,dt_n = \int_0^{t_1} w_1(t_2)\,L_{n-1}(t_1-t_2)\,dt_2,$$

$$g_n(t_1) = \int_0^{t_1} g(t_2)\,dt_2 \int_0^{t_1-t_2} g(t_3)\,dt_3 \ldots \int_0^{t_1-t_2-\ldots-t_{n-1}} g(t_n)\,g(t_1-t_2-\ldots-t_n)\,dt_n.$$

It is obvious that

$$L_0 = 1, \qquad g_0(t_1) = 1,$$

$$L_1(\tau_1) = w_1(\tau), \qquad g_1(t_1) = g(t),$$

$$L_2(\tau_1) = \int_0^{\tau_1} w_1(\tau_2)\, w_1(\tau_1 - \tau_2)\, d\tau_2, \qquad g_2(t_1) = \int_0^{t_1} g(t_2)\, g(t_1 - t_2)\, dt_2,$$

$$L_3(\tau_1) = \int_0^{\tau_1} w_1(\tau_2)\, d\tau_2 \int_0^{\tau_1 - \tau_2} w_1(\tau_3)\, w_1(\tau_1 - \tau_2 - \tau_3)\, d\tau_3,$$

$$g_3(t_1) = \int_0^{t_1} g(t_2)\, dt_2 \int_0^{t_1 - t_2} g(t_3)\, g(t_1 - t_2 - t_3)\, dt_3 \text{ ,etc.}$$

where g (t) is the weight function of linear portion No. 2.

Thus, the system's output, Y (t), will have the following form:

$$Y(t) = m_0 \int_0^T w_0(\tau)\, d\tau + \int_0^T [m'(t-\tau) + n(t-\tau)]\, w_1(\tau)\, d\tau.$$

Then, the error, ϵ (t), in signal reproduction will consist of two parts:

a nonrandom part

$$\varepsilon_0 = H(p)\, m_0 - m_0 \int_0^T w_0(\tau)\, d\tau; \tag{5}$$

and a random part

$$\varepsilon_1(t) = H(p)\, m'(t) - \int_0^T [m'(t-\tau) + n(t-\tau)]\, w_1(\tau)\, d\tau,$$

where H (p) is the transformation operator.

If we introduce the static coefficient, C_0, of dynamic error and use approximately the same reasoning as was followed in [5], we obtain the following expression for C_0 by making use of formula (5):

$$C_0 = H_0 - \int_0^T w_0(\tau)\, d\tau, \tag{6}$$

where $H_0 = H(p)\,|_{p=0}$.

If we substitute the expression for w_0 (t) from (4) in (6), we obtain a condition relating to the weight function w_1 (t):

$$C_0 = H_0 - \frac{1}{q}\left\{\int_0^T w_1(\tau)\, d\tau - \nu \int_0^T d\tau \int_0^\tau g_1(\tau - \theta)\, w_2(\theta)\, d\theta + \ldots \right.$$

$$\left. \ldots + (-1)^{n-1} \nu^{n-1} \int_0^T d\tau \int_0^\tau g_{n-1}(\tau - \theta)\, w_n(\theta)\, d\theta + \ldots \right\}, \tag{7}$$

We shall characterize the random component of the signal reproduction error by its mean square value

$$\overline{\varepsilon_1^2} = \lim_{\theta \to \infty} \frac{1}{2\theta} \int_{-\theta}^{\theta} \left\{ H(p)\, m'(t) - \int_0^T [m'(t-\tau) + n(t-\tau)]\, w_1(\tau)\, d\tau \right\}^2 dt. \tag{8}$$

Thus, the problem of optimizing the system reduces to the finding of a weight function $w_1(t)$ which minimizes expression (8) and which satisfies additional condition (7).

This typical problem of variational calculus leads to the following integral equation

$$\int_0^T [R_m(t-\tau) + R_n(t-\tau)]\, w_1(\tau)\, d\tau = \frac{\gamma_0}{q} \left\{ 1 - 2\nu \int_0^{T-t} L_1(\tau) f_1(T-t-\tau)\, d\tau + \right.$$

$$+ 3\nu^2 \int_0^{T-t} L_2(\tau) f_2(T-t-\tau)\, d\tau + \ldots$$

$$\left. \ldots + (-1)^{n-1} \nu^{n-1} n \int_0^{T-t} L_{n-1}(\tau) f_{n-1}(T-t-\tau)\, d\tau + \ldots \right\} +$$

$$+ \int_{-\infty}^{\infty} R_m(t-\tau)\, \varkappa(\tau)\, d\tau, \tag{9}$$

where γ_0 is a Lagrange multiplier,

$$f_n(u) = \int_0^u g_n(\theta)\, d\theta, \quad \varkappa(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} H(j\omega)\, e^{j\omega t}\, d\omega.$$

Equation (9) can be solved by the method of successive approximations. For this, $w_1(t)$ must be given in the form of a power series in the small parameter ν:

$$w_1(t) = v_0(t) + \nu v_1(t) + \nu^2 v_2(t) + \ldots \tag{10}$$

By substituting (10) in Eqs. (9), we obtain, for each approximation $v_i(t)$, an integral equation of the type

$$\int_0^T [R_m(t-\tau) + R_n(t-\tau)]\, v_i(\tau)\, d\tau = F(t),$$

methods for the solution of which are presented in [6].

By thus determining weight function $w_1(t)$, and then $\Phi_1(p)$, we find, from formulae (1), the sought-for transfer function, $\Phi(p)$, of the linear portion.

2. Practical Design Method

Since q and ν are functions of the mathematical expectation and the dispersion of the signals applied to the inputs of the nonlinear elements, the problem can only be solved by the method of successive approximations.

The following method of solution is suggested.

1. Starting from the conditions of the concrete physical problem, one writes Eq. (9) and finds its solution, limiting oneself to the necessary number of approximations. The expression obtained for $w_1(t)$ will contain the unknown constants A_i which are found by substituting $w_1(t)$ in Eq. (9) and in condition (7). Since the unknown quantities q and ν also enter in (7), the A_i will be functions of these quantities.

2. Starting then with some initial values of q and ν (knowing the type of nonlinear elements, the roughly determined statistical characteristics of the signals applied to the nonlinear element inputs, etc), one computes the constants A_i and, consequently, $w_1(t)$ and $\Phi_1(p)$.

3. By finding concrete expressions for the functions

$$q = \frac{l_1(m_V,\ D_V)}{l_0(m_V,\ D_V)} \text{ and } \nu = k_0(m_y,\ D_y) - k_1(m_y,\ D_y)\, q(m_V,\ D_V)$$

(from [7], for example) and by taking (3) into account, one then finds the values of m_V, D_V, m_Z and D_Z from the following equations:

$$m_V = m_0\left\{1 - \frac{1}{q}\left[\Phi_1(0) - \nu\Psi(0)\,\Phi_1^2(0) + \nu^2\Psi^2(0)\,\Phi_1^3(0) - \ldots\right]\right\},$$

$$D_V = 2\int_0^\infty S_\varphi(\omega)\,|1 - k_1\Psi(j\omega)\,\Phi_1(j\omega)|^2\,d\omega, \tag{11}$$

$$m_Z = \frac{m_0}{q}\left\{\Psi(0)\,\Phi_1(0) - \nu\Psi^2(0)\,\Phi_1^2(0) + \nu^2\Psi^3(0)\,\Phi_1^3(0) - \ldots\right\},$$

$$D_Z = 2\int_0^\infty S_\varphi(\omega)\,|\Psi(j\omega)\,\Phi_1(j\omega)|^2\,d\omega.$$

These equations can be solved graphically.

4. From the values found for m_V, D_V, m_Z and D_Z one computes new values of q and ν and, consequently, new values of the constants A_l, etc.

These computations continue until the latest values of the A_l differ only slightly from the previous set of values.

5. One then finds the final value of $w_1(t)$, then $\Phi_1(p)$ and, from formulae (1), $\Phi(p)$, the sought-for transfer function of the system's linear portion.

3. The Case when the Signal Contains a Regular Function of Time

In the case when the signal contains a regular component, the method described above cannot be given a rigorous basis since the coefficients of the statistical linearization will be functions of time and, consequently, formulae (1) will not, in general, be valid. However, in many practical cases, particularly when the regular component is a slowly varying function of time, one can cite approximate methods of finding the optimal weight function, these methods, however, making no pretense of rigor.

Let $X(t) = a(t) + m(t) + n(t)$ be the input stimulus, where $a(t) + m(t)$ is the signal component. We assume that $a(t)$ has r first bounded derivatives. The assumptions with respect to $m(t)$ and $n(t)$ are as before.

In order to use the method previously presented, it is necessary that, in formulae (1), $k_0(t)$, $k_1(t)$, $l_0(t)$ and $l_1(t)$ be replaced by some average values of them, k_0, k_1, l_0 and l_1. Then, the only quantities subject to change are condition (7), Eq. (9) and the expressions for the mathematical expectations in formulae (11) which, for the given case, take the form:

$$C_l = l!\ H_l + (-1)^{l+1}\frac{1}{q}\left\{\int_0^T \tau^l w_1(\tau)\,d\tau - \nu\int_0^T \tau^l\,d\tau\int_0^\tau g_1(\tau-\theta)\,w_2(\theta)\,d\theta + \ldots\right.$$

$$\left. \ldots + (-1)^{n-1}\nu^{n-1}\int_0^T \tau^l\,d\tau\int_0^\tau g_{n-1}(\tau-\theta)\,w_n(\theta)\,d\theta + \ldots\right\} \quad (l = 0, 1, \ldots, r), \tag{12}$$

$$\int_0^T [R_m(t-\tau) + R_n(t-\tau)]\,w_1(\tau)\,d\tau = \frac{1}{q}\sum_{l=0}^{r}\gamma_l\left\{t^l - 2\nu\int_0^{T-t} L_1(\tau)\,f_1(T, t+\tau)\,d\tau + \right.$$

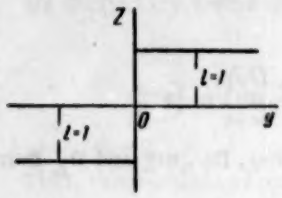

Fig. 2.

$$+ 3\nu^3 \int_0^{T-t} L_2(\tau) f_2(T, t+\tau)\, d\tau - \ldots$$

$$\ldots + (-1)^n n\nu^{n-1} \int_0^{T-t} L_{n-1}(\tau) f_{n-1}(T, t+\tau)\, d\tau + \ldots \Big\} +$$

$$+ \int_{-\infty}^{\infty} R_m(t-\tau)\, x(\tau)\, d\tau, \qquad (13)$$

where $f_n(T, \tau) = \int_0^{T-\tau} (\tau+\theta)^l g_n(\theta)\, d\theta$, γ_l are Lagrange multipliers,

$$m_V(t) = a(t) - \sum_{l=0}^{r} \frac{\Phi_0^{(l)}(0)}{l!} a^{(l)}(t), \quad m_Z(t) = \sum_{l=0}^{r} \frac{\Phi_0^{(l)}(0)}{l!} a^{(l)}(t), \qquad (14)$$

where $\Phi_\theta(p)$ is taken from formula (3).

The expressions for D_V and D_Z in (11) are not changed, since with the given choice of limits of integration from 0 to T, the correlation functions of functions V (t) and Z (t) depend on the difference of their arguments and, consequently, their dispersions remain constant.

The method of solving Eq. (13) is the same as for Eq. (9) since they do not differ in principle from each other.

Since, in (14), the mathematical expectations depend on time, it is suggested that the interval [0, T] be decomposed into a series of discrete values of t, t_1, and that graphical solutions of Eqs. (14) be found for each of these. The resulting functions, q (t) and ν (t) are then averaged, and these average values are substituted in the equations for determining the cofficients A_1 in order to then render them more exact.

All remaining computations are carried out analogously to those for the case of a constant signal.

Example 1. We consider the system which consists of a linear correcting element and one relay element connected in the feedback path, the characteristics of the relay element being shown in Fig. 2.

Let $H(p) = 1$, $m(t) = m_0$, $C_0 = 0$, $R_n(\tau) = \sigma^2 e^{-\alpha|\tau|}$.

Then,

$$S_\varphi(\omega) = S_n(\omega) = \frac{1}{\pi} \frac{\alpha c^2}{\alpha^2 + \omega^2}, \quad q = 1, \quad \nu = k_0 - k_1.$$

If we limit ourselves to the zero'th and first powers of ν, we obtain Eq. (9) in the form

$$\int_0^T R_n(t-\tau)\, w_1(\tau)\, d\tau = \gamma_0 \Big[1 - 2\nu \int_0^{T-t} w_1(\tau)\, d\tau\Big].$$

We substitute the first two terms of expansion (10) in Eq. (9). Then v_0 (t) and v_1 (t) must satisfy the following equations:

$$\int_0^T R_n(t-\tau)\, v_0(\tau)\, d\tau = \gamma_0, \qquad (15)$$

$$\int_0^T R_n(t-\tau)\, v_1(\tau)\, d\tau = -2\gamma_0 \int_0^{T-t} v_0(\tau)\, d\tau. \qquad (16)$$

If we recall the expression for solving the equation of Zadeh and Ragazzini [8], we will write the expression for $v_0(t)$ in the form

$$v_0(t) = A_0 + A_1\delta(t) + A_2\delta(t-T). \tag{17}$$

By substituting (17) in Eq. (15) and requiring that it be satisfied identically, we obtain the equations for the coefficients and the expression for γ_0:

$$A_1 = \frac{A_0}{\alpha}, \qquad A_2 = \frac{A_0}{\alpha}, \qquad \gamma_0 = \frac{2c^2}{\alpha} A_0. \tag{18}$$

Then, $v_0(t)$ will have the form:

$$v_0(t) = A_0\left[1 + \frac{1}{\alpha}\delta(t) + \frac{1}{\alpha}\delta(t-T)\right]. \tag{19}$$

By taking (18) and (19) into account, we can write the right member of Eq. (16) in the following way:

$$-2\gamma_0 \int_0^{T-t} v_0(\tau)\,d\tau = -\frac{4c^2}{\alpha}\left(T + \frac{1}{\alpha}\right)A_0^2 + \frac{4c^2}{\alpha}A_0^2 t = B_0 + B_1 t.$$

Consequently, we obtain $v_1(t)$ in an analogous fashion:

$$v_1(t) = A_3 + A_4 t + A_5\delta(t) + A_6\delta(t-T). \tag{20}$$

By substituting (20) in Eq. (16) and requiring that it be satisfied identically, we obtain the following algebraic equations for the determination of the coefficients A_i:

$$\begin{gathered} \frac{A_3}{\alpha} + \frac{A_4 T}{\alpha} + \frac{A_4}{\alpha^2} - A_6 = 0, \qquad \frac{A_3}{\alpha} - \frac{A_4}{\alpha^2} - A_5 = 0, \\ \frac{2c^2}{\alpha} A_4 = B_1 = \frac{4c^2}{\alpha} A_0^2, \qquad \frac{2c^2}{\alpha} A_3 = B_0 = -\frac{4c^2}{\alpha}\left(T + \frac{1}{\alpha}\right)A_0^2. \end{gathered} \tag{21}$$

If we express all the coefficients from (21) and (18) in terms of A_0, and take into account that

$$w_1(t) = v_0(t) + \nu w_1(t),$$

we obtain the following expression for $w_1(t)$:

$$w_1(t) = A_0\left\{\left[1 - 2\nu A_0\left(T + \frac{1}{\alpha}\right)\right] + 2\nu A_0 t + \left[\frac{1}{\alpha} - \frac{2\nu A_0}{\alpha}\left(T + \frac{2}{\alpha}\right)\right]\delta(t) + \frac{1}{\alpha}\delta(t-T)\right\}. \tag{22}$$

To determine the coefficient A_0, we use condition (7) for the weight function $w_1(t)$, this condition taking the following form in our example:

$$1 = \int_0^T w_1(\tau)\,d\tau - \nu\int_0^T d\tau \int_0^\tau w_1(\tau - \upsilon)\, w_1(\upsilon)\,d\upsilon. \tag{23}$$

By substituting (22) in (23), we obtain a fourth-order algebraic equation for A_0

$$Q_4 A_0^4 + Q_3 A_0^3 + Q_2 A_0^2 + Q_1 A_0 + Q_0 = 0, \tag{24}$$

where

$$Q_4 = \nu^3 T\left(\frac{16}{\alpha^3} + \frac{18T}{\alpha^2} + \frac{20T^2}{3\alpha} + \frac{5}{6}T^3\right), \quad Q_3 = -4\nu^2 T\left(\frac{3}{\alpha^2} + \frac{2T}{\alpha} - \frac{T^2}{3}\right),$$
$$Q_2 = \nu\left(\frac{3}{2}T^2 + \frac{6T}{\alpha} + \frac{4}{\alpha^2}\right), \quad Q_1 = -\left(T + \frac{2}{\alpha}\right), \quad Q_0 = 1.$$

Since we agreed to retain only terms in the zero'th and first powers of ν, we set $Q_4 = Q_3 = 0$ in (24). Then, the expression for A_0 can be written as

$$A_0 = \frac{-Q_1 \pm \sqrt{Q_1^2 - 4Q_2Q_0}}{2Q_2}.$$

For further computations, it is necessary to give numerical values to the Q_i. Let $T = 10$ seconds, $\alpha = 1\ \text{sec}^{-1}$, $\nu = 0.05$. Then,

$$A_0 = 0.0907\ \text{sec}^{-1},$$

$$w_1(t) = 0.0816 + 0.0008t + 0.0807\delta(t) + 0.0907\delta(t-T).$$

For purposes of comparison, we give the value of A_0 for the case when $\nu = 0$:

$$A_0|_{\nu=0} = 0.0833\ \text{sec}^{-1}.$$

In order to verify that the chosen value of ν is correct, we compute the mathematical expectation m_y and the dispersion D_y at the input of the nonlinear element for the weight function $w_1(t)$ obtained.

For our example,

$$m_y = m_0[\Phi_1(0) - \nu(m_y, D_y)\Phi_1^2(0)], \tag{25}$$

$$D_y = \frac{2\alpha c^2}{\pi}\int_0^\infty \frac{1}{\alpha^2+\omega^2}|\Phi_1(j\omega)|^2 d\omega. \tag{26}$$

For simplicity of the calculations, we choose c^2 in (26) in such fashion than $D_y = 1$. Then, by setting $\Phi_1(0) = 1.029$ and $\Phi_1^2(0) = 1.058$ in (25), and by assuming that $m_0 = 1$, we get

$$m_y = 1.029 - \nu(m_y)\,1.058. \tag{27}$$

For the example chosen by us, the dependence of ν on m_y can be expressed in the following way:

$$\nu = k_0 - k_1,\ k_0 = \frac{2}{m_y}\bar{\Phi}(m_y),\ k_1 = \frac{1}{2}\sqrt{1 - 4\bar{\Phi}^2(m_y)} + \frac{1}{\sqrt{2\pi}}e^{-0.5m_y^2},$$

where

$$\bar{\Phi}(u) = \frac{1}{\sqrt{2\pi}}\int_0^u e^{-\frac{t^2}{2}}dt.$$

If we solve Eq. (27) graphically, we can convince ourselves that the chosen value, $\nu = 0.05$, is completely applicable and that further sharpening of this value is senseless since it changes only very slightly the value of coefficient A_0 and, consequently, weight function $w_1(t)$ and transfer function $\Phi_1(p)$.

Thus, the optimal transfer function for the linear portion, $\Phi(p)$, will equal

$$\Phi(p) = \frac{\Phi_1(p)}{1 - 0.63\Phi_1(p)},$$

since the coefficient $k_1 = 0.63$ for the value $\nu = 0.05$ chosen by us.

Example 2. We solve the previous example when the imput signal contains a regular component, $a(t) = m_0 + m_1't$.

By again limiting ourselves to terms of the zero'th and first powers of ν in the expansion of $w_1(t)$, we obtain the equations for $v_0(t)$ and $v_1(t)$:

$$\int_0^T R_n(t-\tau)\,v_0(\tau)\,d\tau = \gamma_0 + \gamma_1 t,$$

$$\int_0^T R_n(t-\tau)\,v_1(\tau)\,d\tau = -2\left\{\gamma_0 \int_0^{T-t} v_0(\tau)\,d\tau + \gamma_1 \int_0^{T-t} (t+\tau)\,v_0(\tau)\,d\tau\right\}, \tag{28}$$

the solutions of which have the forms:

$$v_0(t) = A_0 + A_1 t + A_2\delta(t) + A_3\delta(t-T),$$

$$v_1(t) = A_4 + A_5 t + A_6 t^2 + A_7 t^3 + A_8\delta(t) + A_9\delta(t-T). \tag{29}$$

By substituting (29) in Eqs. (28) and requiring that they be satisfied identically, and also taking conditions (12) into account which, in the given example, take the following forms:

$$\int_0^T w_1(\tau)\,d\tau - \nu\int_0^T d\tau \int_0^\tau w_1(\tau-\theta)\,w_1(\theta)\,d\theta = 1,$$

$$\int_0^T \tau w_1(\tau)\,d\tau - \nu\int_0^T \tau d\tau \int_0^\tau w_1(\tau-\theta)\,w_1(\theta)\,d\theta = C_1,$$

we obtain the algebraic equations for the coefficients A_i:

$$A_2 = \frac{A_0}{\alpha} - \frac{A_1}{\alpha^2},$$

$$A_3 = \frac{A_0}{\alpha} + \frac{A_1}{\alpha}\left(T + \frac{1}{\alpha}\right),$$

$$A_4 = -2\left(T + \frac{1}{\alpha}\right)A_0^2 - 2\left(T^2 - \frac{1}{\alpha^2}\right)A_0A_1 - \frac{2}{3}T^3A_1^2,$$

$$A_5 = 2A_0^2 + 2\left(T - \frac{1}{\alpha}\right)A_0A_1 + \left(\frac{4}{\alpha^2} + T^2\right)A_1^2,$$

$$A_6 = 0, \qquad A_7 = -\frac{A_1^2}{3},$$

$$A_8 = -\frac{2}{\alpha}\left(T + \frac{2}{\alpha}\right)A_0^2 + \frac{2}{\alpha}\left(\frac{2}{\alpha^2} - \frac{T}{\alpha} - T^2\right)A_0A_1 - \frac{1}{\alpha}\left(\frac{2T^3}{3} + \frac{T^2}{\alpha} + \frac{6}{\alpha^2}\right)A_1^2,$$

$$A_9 = \frac{2}{\alpha^2}\left(T + \frac{1}{\alpha}\right)A_1^2$$

$$A_0^2B_0 + A_0A_1B_1 + A_1^2B_2 + A_0B_3 + A_1B_4 - 1 = 0,$$

$$A_0^2E_0 + A_0A_1E_1 + A_1^2E_2 + A_0E_3 + A_1E_4 - C_1 = 0,$$

where

$$B_0 = -\nu\left(\frac{3T^2}{2} + \frac{6T}{\alpha} + \frac{4}{\alpha^2}\right), \qquad E_0 = -\nu 2T\left(\frac{T}{3} + \frac{1}{\alpha}\right),$$

$$B_1 = -\nu\left(\frac{4T^3}{3} + \frac{4T^2}{\alpha} - \frac{2T}{\alpha^2} - \frac{4}{\alpha^3}\right), \qquad E_1 = -\nu T^2\left(\frac{7}{12}T^2 + \frac{4T}{3\alpha} - \frac{2}{\alpha^2}\right),$$

$$B_2 = -\nu\left(\frac{7}{24}T^4 + \frac{2T^3}{3\alpha} - \frac{2T^2}{\alpha^2} - \frac{2T}{\alpha^3} + \frac{4}{\alpha^4}\right), \qquad E_2 = -\nu T\left(\frac{T^4}{10} - \frac{2T^3}{\alpha^2} - \frac{2T}{\alpha^3} - \frac{2}{\alpha^4}\right),$$

$$B_3 = T + \frac{2}{\alpha}, \quad B_4 = T\left(\frac{T}{2} + \frac{1}{\alpha}\right), \qquad E_3 = T\left(\frac{T}{2} + \frac{1}{\alpha}\right), \quad E_4 = T\left(\frac{T^2}{3} + \frac{T}{\alpha} + \frac{1}{\alpha^2}\right).$$

By solving these equations for T = 10 seconds, $\alpha = 1\ \text{sec}^{-1}$, $C_1 = 0.1$ second and $\nu = 0.05$, we obtain the following values for the A_i:

$$A_0 = 0.300\ \text{sec}^{-1}, \qquad A_1 = -0.040\ \text{sec}^{-2},$$

$$A_2 = 0.340, \qquad A_3 = -0.140,$$

$$A_4 = -0.670\ \text{sec}^{-1}, \qquad A_5 = 0.130\ \text{sec}^{-2},$$

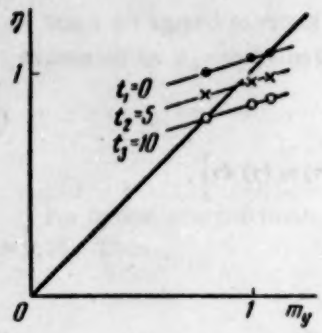

Fig. 3.

$$A_6 = 0, \qquad A_7 = -0.0005 \text{ sec}^{-4},$$
$$A_8 = -0.803, \qquad A_9 = 0.035.$$

Thus, the weight function sought will equal

$$w_1(t) = v_0(t) + \nu v_1(t) = 0.266 - 0.034t + 0.300\delta(t) - 0.138\delta(t-T).$$

By finding

$$\Phi_1(0) = A_0T + A_1\frac{T^2}{2} + A_2 + A_3 = 1.12, \qquad \Phi_1^2(0) = 1.26,$$
$$\Phi_1'(0) = -\left(A_0\frac{T^2}{2} + A_1\frac{T^3}{3} + A_3T\right) = -0.6, \qquad [\Phi_1^{(1)}(0)]^2 = 0.36,$$

we can write Eq. (14), for $m_0 = 1$ and $m_1 = -0.02$, in the form

$$m_y(t) = \eta(t), \quad \eta(t) = (1.132 - \nu\cdot 1.253) - (0.022 - \nu\cdot 0.025)\,t. \tag{30}$$

The dispersion (exactly as in Example 1) we make equal to unity by properly choosing c^2.

By choosing three values of t_i, $t_i = 0$, 5 and 10 seconds, we find the corresponding solutions of Eq. (30) graphically (Fig. 3). For each point of intersection, we obtain $\nu(t_i)$:

$$\nu(0) = 0.03, \qquad \nu(5) = 0.06, \qquad \nu(10) = 0.10.$$

By averaging, we find the new value, $\nu = 0.06$, which must be substituted in conditions (12) to determine the new values of the coefficients A_i. As a result of this computation, we find that

$$A_0 = 0.300 \text{ sec}^{-1}, \quad A_1 = -0.039 \text{ sec}^{-2}, \quad A_2 = 0.339, \quad A_3 = -0.129,$$
$$A_4 = -0.675 \text{ sec}^{-1}, \quad A_5 = 0.127 \text{ sec}^{-2}, \quad A_6 = 0. \quad A_7 = -0.0005 \text{ sec}^{-4},$$
$$A_8 = -0.806, \quad A_9 = 0.033.$$

As is obvious, the values of the coefficients A_i changed very little, so that it is not necessary to sharpen them any further.

Thus, the final value of the weight function $w_1(t)$ is found to be

$$w_1(t) = 0.260 - 0.032t + 0.291\delta(t) - 0.127\delta(t-T).$$

From this we find $\Phi_1(p)$, after which we determine the optimal transfer function of the system's linear portion:

$$\Phi(p) = \frac{\Phi_1(p)}{1 - 0.65\,\Phi_1(p)},$$

where 0.65 is the value of coefficient k_1 for $\nu = 0.06$.

For purposes of comparison, we give the value of $w_1(t)$ for $\nu = 0$ (the problem of Zadeh and Ragazzini):

$$w_1(t) = 0.300 - 0.039t + 0.339\delta(t) - 0.129\delta(t-T).$$

In conclusion, the author wishes to express her gratitude to V. S. Pugachev and Yu. P. Leonov for the aid they have rendered in carrying out the present work.

LITERATURE CITED

[1] L. A. Zadeh, Optimum nonlinear filters, J. Appl. Phys. vol. 24, No. 4 (1953).

[2] P. L Kuznetsov, R. L. Stratanovich and V. I. Techonov, "The passage of random functions through nonlinear systems," [In Russian] Automation and Remote Control (USSR) 16, 3 (1954).

[3] G. C. Newton, Compensation of feedback-control systems subject to saturation. J. Franklin Inst. vol. 254, No. 4, 5 (1952).

[4] L E. Kazakov, "Approximate probabilistic analysis of the accuracy of operation of essentially nonlinear systems," [In Russian]. Automation and Remote Control (USSR) 17, 5 (1956).

[5] V. V. Solodovnikov and P. S. Matveev, "The synthesis of correcting devices for servo systems with noise present with given requirements on dynamic accuracy," [In Russian] Automation and Remote Control (USSR) 16, 3 (1955).

[6] J. H. Laning and R. H. Battin, Random Processes in Automatic Control [Russian translation] IL (1958).

[7] V. S. Pugachev, The Theory of Random Functions and its Applications to Automatic Control Problems [In Russian] Gostekhizdat (1957).

[8] L. A. Zadeh and J. R. Ragazzini, Extension of Wiener's theory of prediction. J. Appl. Phys. vol. 21, No. 7 (1950).

Received May 18, 1959

OPTIMAL CONTROL OF AN OBJECT WITH TWO CONTROLLING STIMULI*

E. A. Rozenman

(Moscow)

The problem considered is that of the optimal transient response in an object at whose input the product of two independent controlling functions acts, namely, an electric motor controlled by the independent variation of the armature current and the excitation current. It is shown that the extremals are combinations of δ-functions.

In recent times a great deal of attention has been given to the question of optimal control of regulated objects and, in particular, to electric motors used as executive organs in various automatic devices. There are works [1-4] devoted to the determination of the most advantageous law of armature current variation for which the duration of the transient response is minimal. With this, the excitation current is assumed to be invariant.

A significant portion of the results of the works cited can be obtained on the basis of the general theory of optimal control, as developed in [5, 6, 7]. In the present paper, as in [8], we consider the problem, of practical interest, of the optimal transient responses in electric motors whose control is implemented by the simultaneous and independent variation of the excitation and the armature currents. This problem is not immediately contained in the general theory mentioned above [5, 6, 7] since it leads to extrema from the class of generalized functions.

Thus, we consider an electric motor wherein, as controlling stimuli, the armature current and the excitation current are used. With the static load and saturation ignored, and with the proper normalization, the equations of motion can be written as

$$\frac{dx}{dt} = y, \qquad \frac{dy}{dt} = uv, \tag{1}$$

where $\underline{x}$ is displacement, $\underline{y}$ is velocity and $\underline{u}$ and $\underline{v}$ are proportional, respectively, to the armature current and the excitation current. We consider the transition from the initial state ($x = x_0$, $y = y_0$) to the given state ($x = x_k$, $y = y_k$).

We shall have the following two problems [8] in mind.

1. To so determine u (t) and v (t) that, for a fixed duration t_0 of the process, the energy dissipation ϑ_0 will be minimal, i.e.,

$$\vartheta_0 = \int_0^{t_0} (u^2 + v^2)\, dt \to \min. \tag{2}$$

*Presented April 15, 1959 at the all-Moscow seminar on nonlinear problems of automatic control theory, held at the IAT AN SSSR.

2. To so determine u (t) and v (t) that the duration of the transient response will be a minimum for a bounded dissipation of energy, i.e., so that $t_0 \to \min$ for

$$\vartheta_0 \leqslant \vartheta_k. \tag{3}$$

Both problems, although having different practical meaning, lead to the same extrema.

We will consider initially the first of these problems. We write the boundary conditions in integral form

$$y_k - y_0 = \int_0^{t_0} uv\,dt, \quad x_k - x_0 - y_0 t_0 = \int_0^{t_0} (t_0 - t)\,uv\,dt. \tag{4}$$

These may be considered as isoperimetric relationships which limit the choice of the function u (t) and v (t).

We now set up the auxiliary functional:

$$J_1 = \int_0^{t_0} [(u^2 + v^2) + \lambda_1 uv + \lambda_2 (t_0 - t)\,uv]\,dt, \tag{5}$$

where λ_1 and λ_2 are constants. On the basis of the Euler equations, we find from (5) that

$$2u + \lambda_1 v + \lambda_2 (t_0 - t)\,v = 0, \quad 2v + \lambda_1 u + \lambda_2 (t_0 - t)\,u = 0. \tag{6}$$

System (6) of equations is homogeneous and has the unique solution u = v = 0 which can not be extremal. We replace the zeroes in the right members of (6) by $\pm \epsilon \to 0$. System (5) can then be satisfied for

$$u = v \neq 0, \tag{7a}$$

if the signs of the right members are the same, and for

$$u = -v \neq 0, \tag{7b}$$

if the signs of the right members are different.

The validity of this is attested to by the fact that the functional ϑ_0 depends in an identical fashion on the controlling functions u (t) and v (t) [Cf., (1) and (2)].

However, the functions u (t) and v (t) may not be defined by (6). It is possible to assume that the extrema, if they exist, lie on the boundaries of the domains of the given functions $\underline{u}$ and $\underline{v}$ and, consequently, go to infinity, since no bounds were placed on the instantaneous values of the controlling functions. Condition (2) requires, however, that the integral ϑ_0 be finite and minimal. If the assumption just made is valid, then u (t) and v (t) can be functions of the δ-function type, i.e., for all values of $t > 0$, u (t) = v (t) = 0 with the exception of one or two points where these functions may take the values $+\infty$ or $-\infty$. We now show that such a law of variation of u (t) and v (t) is in fact optimal.

We consider the simple case of the transition of the system from one equilibrium state to another, i.e., we set x (0) = y (0) = 0 and $y_k = 0$. In this case, the transient response is symmetric with respect to time $t = t_0/2$. We shall seek extremals in the class of binomial functions, and shall set

$$\begin{aligned} u(t) = v(t) = c\left(1 - \frac{2t}{t_0}\right)^m \quad &\text{for} \quad 0 < t < \frac{t_0}{2}, \\ u(t) = -v(t) \quad &\text{for} \quad \frac{t_0}{2} < t < t_0. \end{aligned} \tag{8}$$

By integrating, we find that

$$y(t) = \frac{1}{2}\vartheta(t) = \frac{c^2}{2m+1}\frac{t_0}{2}\left[1-\left(1-\frac{2t}{t_0}\right)^{2m+1}\right],$$

$$x(t) = \frac{c^2}{2m+1}\frac{t_0}{2}\left[t - \frac{t_0}{2(2m+2)}\left(1-\frac{2t}{t_0}\right)^{2m+2} - \frac{t_0}{2(2m+2)}\right]. \tag{9}$$

If we take into account that, for $t = t_0/2$, $x = x_k/2$ and $\vartheta = \vartheta_0/2$, we find that

$$\vartheta_0 = \frac{4c^2}{2m+1}\frac{t_0}{2}, \quad x_k = \frac{c^2}{m+1}\left(\frac{t_0}{2}\right)^2. \tag{10}$$

Fig. 1.

Whence

$$\vartheta_0 = \frac{4x_k}{t_0}\frac{2m+2}{2m+1}. \tag{11}$$

An analogous relationship can be obtained by considering the time interval $t_0/2 < t < t_0$. It is easily shown that $\frac{\partial\vartheta_0}{\partial m} = -\frac{4x_k}{t_0}\frac{2}{(2m+1)^2} < 0$ for any m. Consequently, as $m \to \infty$, the magnitude of ϑ_0 tends to a least value:

$$\lim_{m\to\infty}\vartheta_0 = \frac{4x_k}{t_0} = \vartheta \text{ min}. \tag{12}$$

As $m \to \infty$, the controlling functions u (t) and v (t) reduce to δ-functions. Indeed, for $t = 0$, the binomial in expression (8) equals one, but the initial value of

$$c = \frac{2}{t_0}\sqrt{x_k(m+1)} \to \infty \quad \text{as} \quad m \to \infty.$$

For $t > 0$ and $m \to \infty$, the binomial tends to zero and, by eliminating the indeterminate expression $\infty \cdot 0$, we can convince ourselves that u (t) and v (t) tend to zero.

For the functions y (t) and x (t) we obtain, correspondingly, $y(t) = x_k/t_0 = \text{const}$ and $x(t) = x_k t/t_0$ (Fig. 1).

We can obtain analogously for the interval $t_0/2 < t < t_0$ that, during the entire interval, $u(t) = v(t) = 0$ except for the points $t \to t_0$, where $u(t) \to \infty$, and $v(t) \to \infty$.

We note that, in the example considered, the choice of the class of the functions u (t) and v (t) is without value. We will be led to the same result if we seek the extrema, for example, in the class of exponential functions

$$u(t) = ae^{-mt} \tag{13}$$

or step-functions (Fig. 2). In the latter case, the least value $\vartheta_0 \min$ is attained for $\tau \to 0$ and $a \to \infty$.

The second problem formulated above, namely, the problem of finding the least duration t_0 of the transient response with a bounded value $\vartheta_0 \leq \vartheta$, is related to the minimization of a functional of the following form:

$$J_2 = t_0 + \int_0^{t_0}[\lambda_1(u^2+v^2) + \lambda_2 uv + \lambda_3(t_0-t)uv]\,dt. \tag{14}$$

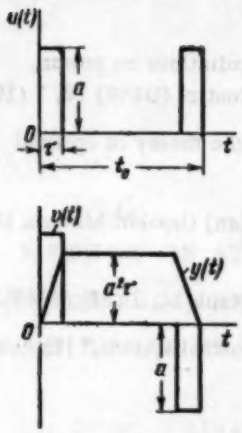

Fig. 2.

Analogously to the previous case, we find that $\partial t_0/\partial m < 0$ for any m, and

$$\lim_{m\to\infty} t_0 = \frac{4x_k}{\vartheta_k} = t_{0\,\min}. \tag{15}$$

The extrema in this case also consist of δ-functions.

The expression

$$t_{0\min}\vartheta_{0\min} = 4x_k \tag{16}$$

relates the least duration of the process with the least dissipation of energy.

The optimal control law obtained can not, obviously, be realized, since an infinite controlling stimulus can not be implemented. Therefore, for the determination of a practically applicable optimal process, there should be imposed on at least one of the controlling functions, in addition to condition (3), a modulus limitation

$$|u| \leqslant u_m. \tag{17}$$

For $u(t) = u_m = \text{const}$, one obtains a linear optimal law of variation for v (t) [Cf. (3), (4)]:

$$v(t) = v_0\left(1 - \frac{t}{t_a}\right). \tag{18}$$

It can be shown that, in the problem considered with the limitation expressed in (17), the optimal transient response is defined by the expressions:

$$\begin{aligned} v(t) &= v_0\left(1-\frac{t}{t_a}\right) && (0 \leqslant t < t_a),\\ u(t) &= u_m = \text{const} \\ v(t) &= u(t) = 0 && (t_a \leqslant t \leqslant t_0 - t_a),\\ v(t) &= -v_0\left(1-\frac{t_0-t}{t_0-t_a}\right) && (t_0-t_a \leqslant t \leqslant t_0).\\ u(t) &= u_m = \text{const} \end{aligned}$$

The magnitude of t_a is so defined that integral (2) will be minimal. The magnitude of v_0 is determined by the boundary conditions.

In the problem considered, both controlling functions, u (t) and v (t), were independent, and the minimization of functional (2) was determined by the choice of these functions. Another formulation of the problem is possible, wherein one of the functions, u (t) for example, is chosen beforehand on the basis of some consideration of another. Then, functional (2) depends only on the one argument v (t), and the extremal is defined by (6):

$$v(t) = -\frac{\lambda_1}{2}u(t) - \frac{\lambda_2}{2}(t_0 - t)\,u(t), \tag{19}$$

and the multipliers λ_1 and λ_2 are defined on the basis of the boundary conditions.

LITERATURE CITED

[1] K. L. Kozhevnikov, "Current diagrams of motors of auxiliary mechanisms of rolling mills," [In Russian] Élektrichestvo 6 (1956).

[2] L. V. Karnyushin, "On rational laws of motion of electric drives with multiple short term modes of operation," [In Russian] Élektrichestvo 6 (1957).

[3] E. A. Rozenman, "On the optimal transient responses in systems with power limitations," [In Russian], Automation and Remote Control (USSR) 18, 6 (1957). •

[4] E. A. Rozenman, "On the limiting speeds of response of servo systems with limitations on power, torque and velocity of the executive elements," [In Russian], Automation and Remote Control (USSR) 19, 7 (1958).•

[5] B. G. Boltyanskii, R. V. Gamkrelidze and L S. Pontryagin, "Contributions to the theory of optimal processes," [In Russian], Doklady AN SSSR 110, 1 (1956).

[6] R. V. Gamkrelidze, "Theory of optimal processes in linear systems," [In Russian] Uspekhi Matem. Nauk 5 (77) (1957).

[7] L. S. Pontryagin, "Optimal control processes," [In Russian], Uspekhi Matem. Nauk 14, 1 (85) (1959).

[8] Sh. Sh. Khamitov, "Investigation of a dc motor as the object of an optimal control system," [In Russian], Elektrichestvo 5 (1958).

Received April 27, 1959

• See English translation .

A METHOD OF SYNTHESIZING LINEAR SAMPLED-DATA AUTOMATIC CONTROL SYSTEMS TO ACCORD WITH DYNAMIC CRITERIA

L. N. Volgin

(Moscow)

A method of synthesizing sampled-data systems is proposed. The transfer function of a system is written as the product of the so-called "minimal polynomial," to which the given portion of the system is reduced, by an artificial portion which is introduced in order to obtain the desired quality. The method presented permits the use of a digital computer for the design of linear sampled-data automatic control systems.

The use of sampled-data correcting filters permits significant improvements in the quality of automatic control systems. The analysis of sampled-data control systems is widely covered in the literature (Cf., for example, [1-5]). Works [6-8] are devoted to the problems of synthesizing optimal sampled-data* systems in accordance with dynamic criteria. Work [6] emphasizes an important factor which must be taken into account in synthesis, namely, the attempt to reduce the zeroes of the desired system's transfer function which are located in the unit circle on the plane of the complex variable $\underline{z}$ disturbs the "roughness" of the system. The existence of such nonreduced zeroes leads to the formation of a "minimal polynomial." The present work gives a method of synthesizing sampled-data servo systems which have stable continuous portions and which possess m'th-order astatism and minimum control times. In [7] there was given a clear presentation of the idea of factoring the polynominals which enter into the transfer function's expression. Here there is presented a purely analytical method of designing a system in accordance with the criterion of minimum mean square error. In [8] the design of sampled-data control systems is considered as a problem in the theory of dynamic programming. The authors, using a matrix representation of the system and the signals, compute the parameters of a linear sampled-data correcting device which provides a minimization of the square quality criteria.

In the present work we present a method which allows one to synthesize optimal linear systems by using the techniques of nonlinear programming. This method permits one to obtain systems which satisfy the "roughness" conditions and the given requirements on process quality.

1. Posing of the Problem

The transfer function $\Phi(z)$ of a single-loop linear sampled-data system is

$$\Phi(z) = \frac{N(z)W(z)}{L(z) + M(z)W(z)}, \qquad (1)$$

where $\underline{z}$ is the lag operator, related to the differentiation operator $\underline{p}$ and the pulse repetition time T by the

*A digital system with a large word-length and operating in accordance with a linear program can also be considered as a linear sampled-data system.

relationship $z = \exp(-pT)$, N (z), L (z) and M (z) are polynomials in $\underline{z}$ with real coefficients which depend on the system parameters, and W (z) is the transfer function of the sampled-data filter, and is the ratio of two polynomials:

$$W(z) = \frac{C(z)}{D(z)}. \tag{2}$$

The sampled-data's transfer function $\Phi(z)$ is a rational-fractional function of $\underline{z}$. The locus of the zeroes and poles of Φ on the $\underline{z}$ plane is a complete characteristic of the system properties.

The introduction of the sampled-data filter in the system changes the number and locus of the initial system's singular points, allows its poles to be taken out to infinity, liquidates its zeroes by reduction and introduces artificial singular points. The interdiction, imposed by the conditions for maintaining system "roughness,"* on reducing the zeroes placed in the circle $|z| = 1$ limits the possibility of completely changing the properties of the original system. The fundamental distinction between singular points lying on the different sides of the unit circle necessitates the factoring of all polynomials which enter into the expression for the system's transfer function. The factoring operation consists of decomposing each polynomial, A, into the product of two factors:

$$A(z) = A^+(z)\,A^-(z), \tag{3}$$

whose roots are distributed on different sides of the unit circle. The factor A^+ is formed of the roots for which $|z| \geq 1$, and A^- of the roots for which $|z| < 1$. In the case of an m-fold root for which $|z| = 1$, one root can be assigned to A^+ and the others to A^-.*

After all the polynomials entering into Φ have been factored, it is necessary to choose the free polynomials in such manner that the system's properties are changed in the right direction. This operation is carried out by the algorithm described in section 2 which follows.

2. Algorithm for the Synthesis of Linear Sampled-Data Automatic Control Systems

The zeroes of the original system's transfer function which cannot be eliminated without disturbing the system's "roughness" form the "minimal polynomial."

Thus, the system's transfer function can be brought to the form

$$\Phi(z) = \Phi_0(z)\,F(z), \tag{4}$$

where $\Phi_0(z)$ is the minimal polynomial and F (z) is the artificial portion, a rational-fractional function

$$F(z) = \frac{A(z)}{B(z)}. \tag{5}$$

The algorithm in accordance with which one chooses the transfer function of the sampled-data filter, W (z), can be written in the form of the following set of rules.

1. We factor the polynomials N (z), L (z) and M (z):

$$N = N^+N^-, \quad L = L^+L^-, \quad M = M^+M^-.$$

2. We decompose the transfer function W into two parts: $W = W_0W^+$ and, with the choice of W^+, we carry out the possible reductions which will not disturb the system's "roughness," and we also introduce an artificial zero, A.

3. By the choice of $W_0 = C_0/D_0$ we reduce the remaining zeroes, and also introduce the artificial pole B. This leads to a polynomial equation which is solvable by some polynomials C_0 and D_0 of minimal degree (cf., the Appendix).

*One can easily convince oneself that reducing (eliminating) a multiple root for which $|z| = 1$ would disturb the system's "roughness."

4. So as not to infringe the condition of physical realizability of the sampled-data filter, $W(0) < \infty$, we fix certain coefficients of the low powers of z in C_θ and D_θ and, in order for the system of equations for C_θ and D_θ to remain consistent, we increase the degrees of polynomials C_θ and D_θ.

5. Starting from the quality criteria, we choose the artificial portion. The quality criteria can either be formulated analytically or given in the form of a system of requirements on the transient response. The method of presenting the quality criteria affects only the method of seeking an optimum which will be used for programming, and does not affect the method of presenting the initial data for setting up the program.

Since, in the general case, the minimal polynomial depends on the artificial portion, one must, to choose the optimal system, solve a joint system of equations for finding the minimal polynomial and estimate.

In the case when the minimal polynomial is independent of the artificial portion, the calculations are essentially simplified.

3. Illustration of the Method as Applied to the Simplest Structures

As illustrations of the algorithm presented, we give the computations for two standard sampled-data automatic control systems.

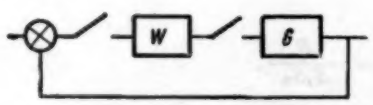

Fig. 1.

System 1. The block schematic is shown in Fig. 1, where $G(z)$ is the system's continuous portion and $W(z)$ is the sampled-data correcting filter.

The system's transfer function has the form:

$$\Phi(z) = \frac{W(z)\,G(z)}{1 + W(z)\,G(z)}.$$

We set $G(z) = P(z)/Q(z)$, where $P(z)$ and $Q(z)$ are polynomials in z.

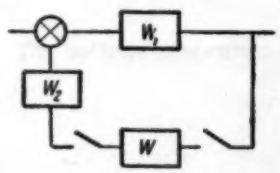

Fig. 2.

We rewrite the system's transfer function in the form

$$\Phi = \frac{W(z)\,P(z)}{Q + W(z)\,P(z)}.$$

The further calculations are carried out in the following order:

1) we decompose P and Q into factors:

$$P = P^+P^-, \quad Q = Q^+Q^-;$$

2) we choose $W^+ = AQ^+/P^+$, and shorten by Q^+/P^+; there remains

$$\Phi = \frac{AP^-C_0}{Q^-D_0 + A^-P^-C_0};$$

3) by equating the denominator to B, we obtain the polynomial equation*

$$Q^-D_0 + AP^-C_0 = B.$$

As the result, we have

$$\Phi = P^-C_0\frac{A}{B}.$$

It follows from this that the minimum polynomial equals

$$\Phi_0 = P^-C_0.$$

*On the solution of the polynomial, cf. the Appendix.

Shortening of C_0 is impossible, since it would lead to the equation

$$Q^- D_0 + (AP^- - B) G_0 = 0,$$

the solution of which is

$$C_0 = Q^-, \; D_0 = AP^- - B.$$

Thus, shortening of C_0 would disturb the system's "roughness."

System 2. The block schematic is shown in Fig. 2, where $W_1(z)$ and $W_2(z)$ are the system's continuous portions and $W(z)$ is the sampled-data correcting filter.

The representation of the transient response has the form:

$$Y = \frac{G(z)}{1 + H(z) W(z)},$$

where

$$G(z) = W_1 X(z) = \frac{P}{QS_1}, \; H(z) = W_1 W_2(z) = \frac{R}{S_1 S_2}.$$

The polynomials S_1, S_2 and Q are the denominators of the z-images of W_1, W_2 and X. After substitution, we obtain

$$Y = \frac{PS_2}{QS_1 S_2 + QRW}.$$

The further computations are carried out in the following order:

1) we carry out the decompositions:

$$P = P^+ P^-, \quad Q = Q^+ Q^-, \quad S_1 = S_1^+ S_1^-, \quad S_2 = S_2^+ S_2^-, \quad R = R^+ R^-;$$

with the condition imposed on the disturbance,

$$Q^- = 1;$$

2) we choose $W^+ = S_2^+/(QR^+A)$; after shortening there remains

$$Y = \frac{PS_2^- AD_0}{RS_1 S_2^- AD_0 + R^- C_0};$$

3) we obtain C_0 and D_0 as solutions of the following polynomial equation:*

$$QS_1 S_2^- AD_0 + R^- C_0 = P^+ B.$$

The representation of the transient response takes the form:

$$Y = P^- S_2^- D_0 \frac{A}{B}.$$

The necessary quality can be obtained by the choice of polynomials A(z) and B(z).

4. Choice of the Artificial Part Providing Optimal Quality

Computing the number and locus of the artificial zeroes and poles which are necessary to attain a given

*On the solution of the polynomial equation, cf. the Appendix.

quality is, in the general case, a complicated task which can only be solved by searching on a digital computer. However, in the simplest cases, the computations can be carried out analytically throughout.

We consider the case when the minimal polynomial is a first-degree polynomial which is independent of the artifical part:*

$$Y_0 = 1 + \alpha z, \ |\alpha| > 1.$$

We introduce one artificial zero:**

$$A = 1 - \sigma z, \ |\sigma| < 1.$$

We obtain the resulting representation of the transient response

$$Y = Y_0 A = (1 + \alpha z)(1 - \sigma z) = 1 + (\alpha - \sigma) z - \alpha \sigma z^2.$$

The criterion

$$J = \sum_{i=0}^{\infty} y_i^2$$

in the given case gives

$$J = 1 + (\alpha - \sigma)^2 + \alpha^2 \sigma^2.$$

The optimal placement of the artificial pole is defined by the condition

$$\frac{\partial J}{\partial \sigma} = 0,$$

from which it follows that, for obtaining an optimum, one must choose

$$\sigma = \frac{\alpha}{1 + \alpha};$$

with which

$$J = 1 + \frac{\alpha^4}{1 + \alpha^2}.$$

The gain in comparison with the case when $\sigma = 0$ is

$$\Delta J = J_0 - J = \frac{\alpha^2}{1 + \alpha^2}$$

which is the greater, the larger is α.

A more powerful means of improving the transient response is the introduction of artificial poles.

For $B = 1 - \sigma z$, we get

$$Y = \frac{Y_0}{B} = \frac{1 + \alpha z}{1 - \sigma z} = 1 + (\alpha - \sigma) z + \sigma (\sigma - \alpha) z^2 + \sigma^2 (\alpha - \sigma) z^3 + \ldots$$

*An example of this case would be System 1 with a stable continuous portion with one zero for $|z| < 1$.

**The introduction of the artificial zero is carried out by choosing the transfer function of the sampled-data filter in accordance with the method presented in Section 3 above.

The estimate by the criterion $J = \sum_{i=0}^{\infty} y_i^2$ equals

$$J = 1 + \frac{(a-\sigma)^2}{1-\sigma^2}.$$

Optimatization gives

$$\sigma = \frac{1}{a}.$$

With this, the gain is unity.

5. Example of the Design of a Sampled-Data Automatic Control System

We consider the system represented in Fig. 2, the transfer functions of whose continuous portions equal:

$$W_1(p) = \frac{p+0.5}{p^2+0.8p+0.25}, \quad W_2(p) = \frac{1-e^{-p}}{p}\frac{1}{p-1}.$$

Let the disturbance applied to the system have the form of a jump, i.e., X (p) = 1/p. W_1 (p) corresponds to the stable oscillations of the system. The poles of W_1 (p) are:

$$p_1 = -0.4 + j\cdot 0.3, \quad p_2 = -0.4 - j\cdot 0.3.$$

The function W_2 (p) contains fixing and unstable elements.

By setting up the functions C (z) and H (z) and making the transition to the variable $\underline{z}$ (by well-known rules, presented in [5] for example), we get

$$G = \frac{-2+1.4z}{(1-z)(1-1.28z+0.45z^2)},$$
$$H = \frac{0.6z(1-0.64z)(1+1,3z)}{(1-2.7z)(1-1.28z+0.45z^2)}.$$

We carry out the factorings:

$$P^+ = -2+1.4z, \qquad P^- = 1.$$
$$Q^+ = 1-z, \qquad Q^- = 1.$$
$$R^+ = 1-0.64z, \qquad R^- = 0.6z(1+1.3z),$$
$$S_1^+ = 1-1.28z+0.45z^2, \qquad S_1^- = 1.$$
$$S_2^+ = 1. \qquad S_2^- = 1-2.7z.$$

To improve the transient response, we introduce one artificial pole:

$$B = 1-\sigma z, \quad A = 1, \quad |\sigma| < 1.$$

We choose

$$W^+ = \frac{1}{(1-z)(1-0.64z)}.$$

The polynomial equation has the form:

$$LD_0 + MC_0 = N,$$

where

$$L = 1 - 5z + 3.1z^2 + 1.3z^3 + 1.1z^4,$$
$$M = 0.6z + 0\ 78z^2,$$
$$N = -2 + (1.4 + 2c)z - 1.4cz^2.$$

We write C_0 and D_0 in the form

$$C_0 = c_0 + c_1 z + c_2 z^2 + c_3 z^3,$$
$$D_0 = d_0 + d_1 z.$$

By expanding the polynomial equation in an algebraic system and introducing the homogeneous notation

$$x_1 = d_1, \quad x_2 = c, \quad x_3 = c_0, \quad x_4 = c_1, \quad x_5 = c_2, \quad x_6 = c_3,$$

we obtain the system of limitations

$$x_1 - 2x_2 + 0.6x_3 = -8.6,$$
$$5x_1 + 1.4x_2 + 0.78x_3 + 0.6x_4 = 6.2,$$
$$3.1x_1 + 0.78x_4 + 0.6x_5 = 2.6,$$
$$1.3x_1 + 0.78x_5 + 0.6x_6 = 2.16,$$
$$1.1x_1 + 0.78x_6 = 0,$$
$$-1 < x_2 < 1.$$

The representation of the transient response has the form:

$$Y = \frac{-2 + (5.4 + x_1)z - 2.7x_1 z^2}{1 - x_2 z}.$$

We take $J = \sum_{i=0}^{\infty} y_i^2$. as the quality criterion. We compute this estimate:

$$J = \frac{32.2 + 10.8x_1 - 21.6x_2 + 8.3x_1^2 - 33.2x_1x_2 + 10.8x_1x_2^2 - 5.4x_1^2x_2}{1 - x_2^2}.$$

The technique of nonlinear programming [9] can be used to find the minimum of this function given the limitations imposed.

APPENDIX

A polynomial equation of the form

$$LD_0 + MC_0 = N$$

is expanded in a system of algebraic equations

$$l_0d_0 + m_0c_0 = n_0$$
$$l_1d_0 + l_0d_1 + m_1c_0 + m_0c_1 = n_1$$
$$l_2d_0 + l_1d_1 + l_0d_2 + m_2c_0 + m_1c_1 + m_0c_2 = n_2,$$
$$\cdots$$

which has a solution if the following conditions hold

$$\delta \geq \mu - 1, \quad \varkappa \geq \lambda - 1, \quad \nu \leq \lambda + \delta = \mu + \varkappa,$$

where λ, μ, δ, $\varkappa$, ν are the degrees of the polynomials, respectively, L, M, D_0, C_0 and N.

The condition of physical realizability of the sampled-data filter requires that $d_0 \neq 0$. If $m_0 = 0$, it is then necessary to choose N such that $n_0 \neq 0$.

The polynomial equation is equivalent to the congruence $LD_0 \equiv N \pmod{N}$, and Euclid's algorithm can also be used for its solution.

LITERATURE CITED

[1] Automatic Control Fundamentals, [In Russian], (edited by V. V. Solodovnikov) Mashgiz (1954).

[2] Ya. Z. Tsypkin, Transient Responses and Steady-State Processes in Sampled-Data Circuits [In Russian], GEI (1951).

[3] Ya. Z. Tsypkin, Theory of Sampled-Data Systems [In Russian] (Moscow, 1958).

[4] H. M. James, N. B. Nichols and R. S. Phillips, Theory of Servomechanisms [Russian translation] ILI (1951).

[5] J. R. Ragazzini and L. A. Zadeh, The Analysis of Pulse Systems, Trans. AIEE, vol. 71, pt. II (1952).

[6] A. R. Bergen and J. R. Ragazzini, Sampled-data processing techniques for feedback control systems. Trans. AIEE, vol. 73, pt. II (1954).

[7] S. S. L. Chang, Statistical Design Theory for Digital-Controlled Continuous Systems. Trans. AIEE, vol. 77, pt. II (1958).

[8] R. E. Kalman and R. W. Koepcke, Optimal synthesis of linear sampling control systems using generalized performance indexes. Trans. ASME, vol. 80, No. 8 (1958).

[9] A. W. Tucker, Linear and nonlinear programming. Operations research, vol. 5, No. 2 (1957).

Received January 26, 1959

AC SERVO SYSTEMS IN WHICH THE SIGNAL DEPENDS ON THE ERROR AND ITS DERIVATIVE

N. P. Vlasov

(Gor'kii)

AC systems are considered. Under the assumption that the signal depends on the error and on the speed with which the error changes, we find the transmission function for a linear four-terminal network and for servo systems which are stabilized by differentiating four-terminal networks and velocity feedback.

In work [1] it was assumed that the function at the output of the modulating device for a servo system can be presented in the form

$$v(t) = x(t)\frac{\partial U_m}{\partial x}\cos(\omega_0 t + \varphi). \tag{1}$$

Here,

$$x(t) = \theta_1(t) - \theta_2(t) \tag{2}$$

is the error and $\theta_1(t)$ and $\theta_2(t)$ are the functions at the input and output respectively of the servo system. The function v (t) is the oscillations of the carrier frequency ω_0 modulated by the error.

Actually, at the output of a modulating device which is implemented, for example, by means of two selsyns, with the assumption that the following conditions hold

$$\dot{\theta}_1(t) \ll \omega_0, \quad \dot{\theta}_2(t) \ll \omega_0. \tag{3}$$

there appears, as can be shown [2], the function

$$u_x(t) = \frac{\partial U_m}{\partial x}\left[x(t)\cos(\omega_0 t + \varphi) + \frac{\dot{x}}{\omega_0}\sin(\omega_0 t + \varphi)\right]. \tag{4}$$

In the present paper, on the basis of the methods developed in [1], we consider servo systems in which the transmission function is specified for the four-terminal network connected in the control loop, as well as systems which are stabilized by a velocity feedback, with the condition that function (4) appears at the modulating device's output.

1. We consider initially a linear four-terminal network V with transmission function Y (p), connected between modulator M and ideal phase detector D [1] (Fig. 1). The Laplace transform of the function u_x (t) will be

$$\overline{u_x(p)} = \frac{1}{2}\frac{\partial U_m}{\partial x}\frac{p}{j\omega_0}\left[\overline{x(p - j\omega_0)}\,e^{j\varphi} - \overline{x(p + j\omega_0)}\,e^{-j\varphi}\right]. \tag{5}$$

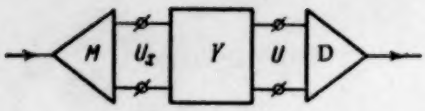

Fig. 1.

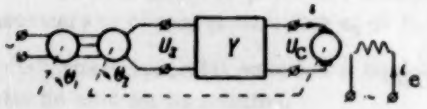

Fig. 2.

The transform of the function u (t) at the phase detector's input is

$$\overline{u(p)} = \overline{u_x(p)\,Y(p)}. \tag{6}$$

Let the phase detector be ideal, and let it implement the multiplication of the functions u (t) and cos ($\omega_0 t + \psi$). Consequently, at the output of the phase detector we shall have

$$u_D(t) = u(t)\cos(\omega_0 t + \psi). \tag{7}$$

The transform of function (7) is

$$\overline{u_D(p)} = \frac{1}{2}\left[\overline{u(p - j\omega_0)}\,e^{j\psi} + \overline{u(p + j\omega_0)}\,e^{-j\psi}\right]. \tag{8}$$

By substituting the value of u (p) from (6) in this expression we get

$$\begin{aligned}\overline{u_D(p)} = \frac{1}{4}\frac{\partial U_m}{\partial x}\Bigg\{\Bigg[&\frac{(p + j\omega_0)}{j\omega_0}\,\overline{Y(p + j\omega_0)}\,e^{j(\varphi-\psi)} - \\ &- \frac{(p - j\omega_0)}{j\omega_0}\,\overline{Y(p - j\omega_0)}\,e^{-j(\varphi-\psi)}\Bigg]\overline{x(p)} + \\ &+ \frac{(p - j\omega_0)}{j\omega_0}\,\overline{Y(p - j\omega_0)}\,e^{j(\varphi-\psi)}\,\overline{x(p - j\cdot 2\omega_0)} - \\ &- \frac{(p + j\omega_0)}{j\omega_0}\,\overline{Y(p + j\omega_0)}\,e^{-j(\varphi+\psi)}\,\overline{x(p + j\cdot 2\omega_0)}\Bigg\}.\end{aligned}$$

If we discard from this expression those terms which correspond to oscillations modulated at twice the carrier frequency, which will not be passed by the low-frequency filter [1], we obtain the following transmission function for the chain of links between the input link and the phase detector:

$$\begin{aligned}\overline{Y_D(p)} = \frac{\overline{u_D(p)}}{\overline{x(p)}} = \frac{1}{4}\frac{\partial U_m}{\partial x}\Bigg[&\frac{(p + j\omega_0)}{j\omega_0}\,\overline{Y(p + j\omega_0)}\,e^{j(\varphi-\psi)} - \\ &- \frac{(p - j\omega_0)}{j\omega_0}\,\overline{Y(p - j\omega_0)}\,e^{-j(\varphi-\psi)}\Bigg].\end{aligned} \tag{9}$$

If the four-terminal network is inertialess and Y (p) = k_0 then, for Y_D (p) we shall have

$$\overline{Y_D(p)} = \frac{k_0}{4}\frac{\partial U_m}{\partial x}\left[\cos(\varphi - \psi) + \frac{p}{\omega_0}\sin(\varphi - \psi)\right]. \tag{10}$$

Thus, by selecting the phase angle ψ of the reference voltage applied to the phase detector, one can not only affect the gain, but can also obtain a component which is proportional to the velocity of the error.

2. We now consider an ac servo system with a single-phase commutator motor as the servomotor, controlled on the armature side.

In this case, the phase detector is the servomotor and, for the transform of the function at the phase detector's input, i.e., for the voltage on the servomotor's armature terminals, we take expression (6) (Fig. 2).

With certain simplifying assumptions, formulated in [1], the servomotor equations can be written as:

the equation of the armature circuit

$$u(t) = u_c(t) = ir + L\frac{di}{dt} + Ci_e\frac{d\theta_2}{dt}; \tag{11}$$

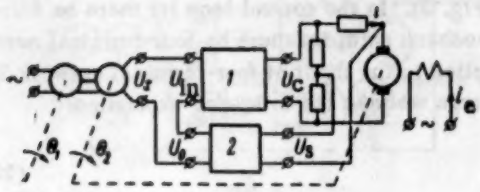

Fig. 3.

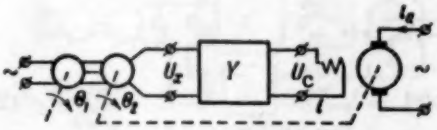

Fig. 4.

the equation of armature rotation

$$J\frac{d^2\theta_2}{dt^2} = Ci_c i - M_c = Ci_c i - k\frac{d\theta_2}{dt}. \qquad (12)$$

The nomenclature used in Eqs. (11) and (12) was introduced in [1].

For the current in the servomotor's excitation winding we take the expression

$$i_c = I_m \cos(\omega_0 t + \psi). \qquad (13)$$

We now take the transforms of Eqs. (11) and (12), taking Eq. (13) into account. We find $\overline{\theta_2(p)}$ from the system obtained, taking Eqs. (5) and (6) into account, having first discarded the terms corresponding to oscillations modulating twice the carrier frequency. The result is

$$\overline{\theta_2(p)} = \frac{\frac{1}{4}\frac{CI_m}{j\omega_0}\frac{\partial U_m}{\partial x}\left\{\frac{(p+j\omega_0)\,\overline{Y(p+j\omega_0)}}{|r+L(p+j\omega_0)|}e^{-j(\psi-\varphi)} - \frac{(p-j\omega_0)\,\overline{Y(p-j\omega_0)}}{|r+L(p-j\omega_0)|}e^{j(\psi-\varphi)}\right\}\overline{x(p)}}{Jp^3 + \left\{k + \frac{(CI_m)^2}{2}\frac{(r+Lp)}{[(r+Lp)^2+(L\omega_0)^2]}\right\}p}.$$

If we go to the argument $p/\omega_0 = s$, we will have, for the transmission function of the open-loop system

$$\frac{\overline{\theta_2(s)}}{\overline{x(s)}} = \frac{1}{2}\frac{\{j(s-j)[T_a(s+j)+1]\,\overline{Y(s-j)}e^{j\delta} - j(s+j)[T_a(s-j)+1]\,\overline{Y(s+j)}\,e^{-j\delta}\}}{T_k s\{k_m(T_m s+1)[(T_a s+1)^2+T_a^2]+T_a s+1\}}. \qquad (14)$$

Here, $\delta = \psi - \varphi$. The values of the constants T_a, T_m, k_m and T_k are given in [1].

We now consider the systems which is stabilized by an ac differentiating four-terminal network [3, 4] with transmission function

$$\overline{Y(s)} = k_0\overline{Y_1(s)}, \qquad (15)$$

where

$$\overline{Y_1(s)} = \frac{s^2+k_1 s+1}{s^2+k_2 s+1}. \qquad (16)$$

For the transmission function of the open-loop system we get

$$\frac{\overline{\theta_2(s)}}{\overline{x(s)}} = k_0\frac{\overline{A(s)}\{\cos\delta - [T_a(s^2+1)+s]\sin\delta\} + \{[T_a(s^2+1)+s]\cos\delta + \sin\delta\}(k_2-k_1)s^2}{T_k s\{k_m(T_m s+1)[(T_a s+1)^2+T_a^2]+T_a s+1\}\,\overline{B(s)}} \qquad (17)$$

Here,

$$\overline{A(s)} = s^2(s^2+4) + (k_1+k_2)(s^2+2)s + k_1k_2(s^2+1), \qquad (18)$$

$$\overline{B(s)} = 2[s^2(s^2+4) + 2k_2(s^2+2)s + k_2^2(s^2+1)]. \qquad (19)$$

We now consider a system stabilized by velocity feedback (Fig. 3). In the control loop let there be four-terminal network 1, with transmission function $\overline{Y(p)}$, and in the feedback path, let there be four-terminal network 2, with transmission function $\overline{Y_2(p)}$. The voltage $u_s(t)$ is applied to the input of four-terminal network 2 from a bridge circuit in the motor's armature circuit. For this voltage we take the following expression:

$$u_s(t) = k_s C i_e \frac{d\theta_2}{dt}. \tag{20}$$

The transform, $\overline{u_s(p)}$, will be

$$\overline{u_s(p)} = \frac{1}{2} k_s C I_m [(p - j\omega_0)\overline{\theta_2(p - j\omega_0)}\, e^{j\varphi} + \\ + (p + j\omega_0)\overline{\theta_2(p + j\omega_0)}\, e^{-j\varphi}]. \tag{21}$$

The transform, $\overline{u_0(p)}$, of the voltage at the output of four-terminal network 2 will be

$$\overline{u_0(p)} = \overline{u_s(p)}\,\overline{Y_2(p)}. \tag{22}$$

Since the transform of the voltage $u_x(t)$ at the differential output is defined by expression (5), the transform of the voltage at the input of four-terminal network 1 will be

$$\overline{u_{in}(p)} = \overline{u_x(p)} - \overline{u_0(p)}. \tag{23}$$

Here, $\overline{u_0(p)}$ is given by Eq. (22).

In this case, the transform of the voltage on the motor's armature winding will be

$$\overline{u_c(p)} = \overline{u_{in}(p)}\,\overline{Y(p)}. \tag{24}$$

By again using Eqs. (11) and (12), taking the transforms of these equations, and also by taking Eqs. (5) and (21)-(24) into account, we obtain the following transmission function of the open-loop servo system, written for the argument s:

$$\frac{\overline{\theta_2(s)}}{\overline{x(s)}} = \frac{1}{2} \frac{j(s-j)[T_a(s+j)+1]\,\overline{Y(s-j)}\,e^{j\delta} - j(s+j)[T_a(s-j)+1]\,\overline{Y(s+j)}\,e^{-j\delta}}{T_k s\{k_m(T_m s+1)[(T_a s+1)^2 + T_a^2] + T_a s + 1 + k_c \overline{H(s)}\}}. \tag{25}$$

Here,

$$\overline{H(s)} = \frac{1}{2}\{\overline{Y_2(s-j)}\,\overline{Y(s-j)}[T_a(s+j)+1] + \\ + \overline{Y_2(s+j)}\,\overline{Y(s+j)}[T_a(s-j)+1]\}. \tag{26}$$

3. We now consider a servo system with a single-phase commutator motor, controlled on the side of the excitation winding (Fig. 4). The phase detector is once again the servomotor, but here expression (6) is the transform of the voltage supplied to the excitation winding. With the simplifying assumptions made in [1, 5], the servomotor equations can be written in the following way:

the equation for the excitation winding circuit

$$u_c(t) = ir + L\frac{di}{dt}; \tag{27}$$

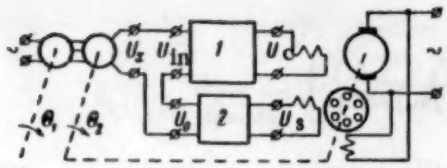

Fig. 5.

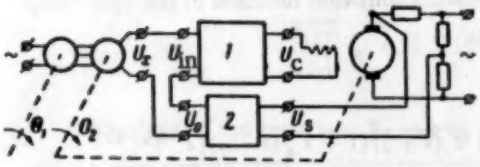

Fig. 6.

the equation for the armature circuit

$$u(t) = i_a r_a + L_a \frac{di_a}{dt} + Ci \frac{d\theta_2}{dt}; \qquad (28)$$

the equation for the armature rotation

$$J \frac{d^2\theta_2}{dt^2} = Cii_a - M_c = Cii_a - k \frac{d\theta_2}{dt}. \qquad (29)$$

The nomenclature used in Eqs. (27)-(29) is given in [1].

We make the assumption that the last term in the right member of Eq. (28) is small; this condition holds for sufficiently small servomotor rotational speeds. Then, with a sinusoidal voltage u (t) applied to the armature terminals, the armature current i_a will also be sinusoidal.

We set

$$i_a = I_m \cos(\omega_0 t + \phi). \qquad (30)$$

We now take the transforms of Eqs. (27) and (29), keeping Eq. (30) in mind; we find $\overline{\theta_2(p)}$ from the system obtained with Eqs. (5) and (6) taken into account, having first discarded the terms corresponding to oscillations modulating the doubled carrier frequency. The result is

$$\overline{\theta_2(p)} = \frac{CI_m \frac{\partial U_m}{\partial x} \left\{ \frac{j(p-j\omega_0)\overline{Y(p-j\omega_0)}}{|r+L(p-j\omega_0)|} e^{j\delta} - \frac{j(p+j\omega_0)\overline{Y(p+j\omega_0)}}{|r+L(p+j\omega_0)|} e^{-j\delta} \right\} \overline{x(p)}}{4\omega_0 (Jp^2 + kp)}.$$

By going to the argument $s = p/\omega_0$, we will have, for the transmission function of the open-loop system,

$$\frac{\overline{\theta_2(s)}}{\overline{x(s)}} = \frac{1}{2} \frac{\{j(s-j)[T_a(s+j)+1]\overline{Y(s-j)}\, e^{j\delta} - j(s+j)[T_a(s-j)+1]\overline{Y(s+j)}\, e^{-j\delta}\}}{K_m T_k s (T_m s+1)[(T_a s+1)^2 + T_a^2]}.$$

Here, the constants T_a, T_m, k_m, T_k and δ have the same values as in Eq. (14).

We next consider a system stabilized by a differentiating four-terminal network with transmission function (15). For the transmission function of the open-loop system we obtain

$$\frac{\overline{\theta_2(s)}}{\overline{x(s)}} = k_0 \frac{\overline{A(s)}\cos\delta - [T_a(s^2+1)+s]\sin\delta + \{[T_a(s^2+1)+s]\cos\delta + \sin\delta\}(k_2-k_1)s^2}{k_m T_k s (T_m s+1)[(T_a s+1)^2 + T_a^2]\overline{B(s)}}.$$

Here, the functions $\overline{A(s)}$ and $\overline{B(s)}$ are defined by expressions (18) and (19). As an example of a system stabilized by a velocity feedback, we consider a system with an asynchronous tachometer generator (Fig. 5). It can be shown that, if a bridge scheme is used for velocity feedback (Fig. 6), the system will be described by a set of nonlinear differential equations with periodic coefficients, for which the method developed in [1] is inapplicable.

For the voltage taken off the tachometer generator, $u_s(t)$, we take the expression

$$u_s(t) = k_c \frac{U_m}{\omega_0} \cos(\omega_0 t + \alpha) \frac{d\theta_2}{dt}.$$

For the transform, $\overline{u_2(p)}$, we obtain

$$\overline{u_2(p)} = \frac{k_c}{2\omega_0} U_m [(p - j\omega_0)\,\overline{\theta_2(p - j\omega_0)}\, e^{j\alpha} + (p + j\omega_0)\,\overline{\theta_2(p + j\omega_0)}\, e^{-j\alpha}]. \quad (31)$$

By again using Eqs. (27)-(30), taking their transforms, and by also taking Eqs. (5), (31) and (22)-(24) into account, we obtain the following expression in the argument $\underline{s}$ for the transmission function of the open-loop system:

$$\frac{\overline{\theta_2(s)}}{\overline{x(s)}} = \frac{j(s-j)\,\overline{Y(s-j)}\,[1 + T_a(s+j)]\,e^{j(\psi-\varphi)} - j(s+j)\,\overline{Y(s+j)}\,[1 + T_a(s-j)]\,e^{-j(\psi-\varphi)}}{T_2 s\,\{(T_1 s + 1)\,[(1 + T_a s)^2 + T_a^2] + T_3 \overline{E(s)}\}}.$$

Here,

$$\overline{E(s)} = \frac{1}{2}\{\overline{Y(s-j)}\,\overline{Y_2(s-j)}\,[1 + T_a(s+j)]\,e^{j(\psi-\alpha)} + \overline{Y(s+j)}\,\overline{Y_2(s+j)}\,[1 + T_a(s-j)]\,e^{-j(\psi-\alpha)}\}.$$

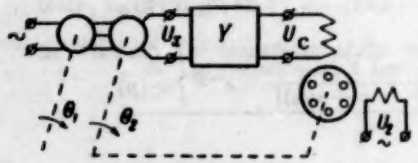

Fig. 7.

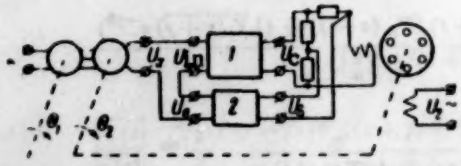

Fig. 8.

The values of the constants T_1, T_2 and T_3 are given in [1].

4. We now consider a servo system with a two-phase asynchronous motor as the phase detector (Fig. 7). With certain simplifying assumptions, formulated in [1], we can write the servomotor's equations of motion in the following way:

the equation of rotor rotation

$$J\frac{d^2\theta_2}{dt^2} = M(t) - M_s(t) = \frac{1}{r}\left[\psi_a u_2 - \psi_e u_c - \frac{d\theta_2}{dt}(\psi_s^2 + \psi_a^2)\right] - k\frac{d\theta_2}{dt}; \quad (32)$$

the equation of the control winding

$$u_c(t) = \frac{d\psi_a}{dt}; \quad (33)$$

the equation of the second stator winding

$$\frac{d\psi_e}{dt} = u_2(t) = U_{2m}\sin(\omega_0 t + \psi). \quad (34)$$

The nomenclature used in Eqs. (31)-(34) was given in [1].

By assuming that the amplitude of the controlling voltage is small compared with the amplitude of voltage u_2 and, consequently, that $\psi_a^2 \ll \psi_e^2$, we discard the term ψ_a^2 from Eq. (31). We now transform Eqs. (31)-(33), find $\overline{\theta_2(p)}$ from the system obtained with Eqs. (5) and (6) being taken into account, having first discarded the terms corresponding to oscillations modulating the doubled carrier frequency. The result is

$$\overline{\theta_2(p)} = \frac{U_{2m}\partial U_m}{4r\omega_0\,\partial x}\,\frac{\left[j\left(\frac{p}{\omega_0}-j2\right)\overline{Y(p-j\omega_0)}e^{j(\psi-\varphi)} - j\left(\frac{p}{\omega_0}+j2\right)\overline{Y(p+j\omega_0)}\,e^{-j(\psi-\varphi)}\right]\overline{x(p)}}{\left(Jp^2+kp+\frac{1}{2r}\,\frac{U_m^2}{\omega_0^2}\,p\right)}.$$

By using the argument $s = p/\omega_0$, we obtain the transmission function of the open-loop system

$$\frac{\overline{\theta_2(s)}}{\overline{x(s)}} = \frac{j(s-j2)\,\overline{Y(s-j)}\,e^{j\delta} - j(s+j2)\,\overline{Y(s+j)}\,e^{-j\delta}}{T_2 s(T_1 s+1)}.$$

The values of the constants T_1 and T_2 are given in [1].

We now consider a servo system stabilized by a differentiating four-terminal network with transmission function (15). In this case, the transmission function of the open-loop system is

$$\frac{\overline{\theta_2(s)}}{\overline{x(s)}} = k_0\,\frac{(2+js)[\overline{A(s)}+j(k_1-k_2)s^2]e^{j\delta} + (2-js)[\overline{A(s)}-j(k_1-k_2)s^2]e^{-j\delta}}{T_2 s(T_1 s+1)\,\overline{B(s)}}.$$

Here, functions $\overline{A(s)}$ and $\overline{B(s)}$ are defined by expressions (18) and (19).

We now consider the servo system stabilized by a velocity feedback (Fig. 8). In the control loop is four-terminal network 1 with transmission function $\overline{Y(p)}$ and, in the feedback path, four-terminal network 2 with transmission function $\overline{Y_2(p)}$. Voltage $u_s(t)$ is applied to the input of four-terminal network 2 from the bridge scheme in the control winding circuit. It can be shown that voltage $u_s(t)$ is proportional to the rotational speed of the motor's rotor, or to voltage $u_2(t)$, but shifted in phase by $\pi/2$ with respect to voltage $u_2(t)$. We take the following expression for voltage $u_s(t)$:

$$u_s(t) = \frac{k_s}{\omega_0}U_{2m}\cos(\omega_0 t+\phi)\,\frac{d\theta_2}{dt}. \tag{35}$$

The transform of $u_s(t)$ is

$$\overline{u_s(p)} = \frac{1}{2}\,\frac{k_s}{\omega_0}U_{2m}[(p-j\omega_0)\,\overline{\theta_2(p-j\omega_0)}\,e^{j\psi} + (p+j\omega_0)\,\overline{\theta_2(p+j\omega_0)}\,e^{-j\psi}]. \tag{36}$$

To find the transform, $\overline{u_c(p)}$, at the input of the control winding, we use formulae (22)-(24).

Taking the transforms of Eqs. (32)-(34) and taking Eqs. (5), (36) and (22)-(24) into account, we obtain the following expression in the argument s for the transmission function of the open-loop system:

$$\frac{\overline{\theta_2(s)}}{\overline{x(s)}} = \frac{[j(s-j\cdot 2)\,\overline{Y(s-j)}\,e^{j\delta} - j(s+j\cdot 2)\,\overline{Y(s+j)}\,e^{-j\delta}](s^2+1)}{T_2 s[(T_1 s+1)(s^2+1)+k_a k_c\overline{E(s)}]}.$$

Here,

$$\overline{E(s)} = \frac{1}{2}[(s-j\cdot 2)(s+j)\,\overline{Y(s-j)}\,\overline{Y_2(s-j)} + (s+j\cdot 2)(s-j)\,\overline{Y(s+j)}\,\overline{Y_2(s+j)}]$$

and

$$k_a = \frac{U_{2m}^2}{2rk\omega_0+U_{2m}^2}.$$

Thus, the methods presented in [1] can be used to obtain the transfer functions for ac systems whose signals depend on the error and on the velocity of the error's changes.

In this paper we obtained the transfer functions for linear four-terminal networks and for servo systems stabilized by differentiating four-terminal networks or by velocity feedback. For this, we considered as phase detectors both single-phase commutator motors (with control on the side of the armature or on the side of the excitation winding) and two-phase asynchronous motors.

LITERATURE CITED

[1] N. P. Vlasov, "Elements of the theory of ac servo systems, and mixed systems" [In Russian]. Izv. Vysshikh Uchebnikh Zavedenii, Ser. Élektromekhanika, 1 (Novocherkassk, 1958).

[2] D. Morris, "Theoretical and experimental methods of analyzing the envelopes of modulated signals in ac servo systems," Automatic Control. Collected Materials of the Conference in Crenfield [Russian translation] (1951).

[3] H. M. James, N. B. Nichols and R. S. Phillips, Theory of Servomechanism [Russian translation] IL (1951).

[4] Collection, Correcting Circuits in Automata [Russian translation of unknown original], pp. 488-490. IL (Moscow, 1954).

[5] E. I. Chernov, "Investigation of low-power ac servomechanisms," [In Russian]. Automation and Remote Control (USSR) 13, 1 (1952).

Received October 8, 1958

GRAPHICOANALYTIC METHOD OF COMPUTING STEADY-STATE SPEED OF PNEUMATIC DRIVES

V. N. Dmitriev and V. I. Chernyshev

(Moscow)

A method is considered for computing the steady-state motion of a pneumatic piston executive mechanism. A comparison is given of the computed and the experimental characteristics of such a mechanism.

Pneumatic pistons are widely used as the executive mechanisms in pneumoautomatic systems. However, very few works, either domestic or foreign, in the literature are devoted to the theory and design of this type of mechanism. Because of the compressibility of air and the variability of the air which is charged into, and discharged from, the cylinder voids in the executive mechanism, the motion of this mechanism is described by a system of complicated nonlinear differential equations. The majority of the authors therefore consider this problem in a linear approximation [1-3]. There are also works [4, 5] which consider the motion of a pneumatic piston executive mechanism without feedback by stages. Such a solution to the problem is complicated.

The method of designing pneumatic piston executive mechanisms which is presented in the present work does not pretend to applicability to an arbitrarily chosen mechanism of this type. However, it can be successfully applied for pneumatic piston executive mechanisms whose pistons have a significant distance of travel, which do not thereby develop large accelerations, and whose load is constant. The class of such systems in industry is quite large. This is a different type of drive than drives for clamping devices, drives for feed mechanisms and terminal organs of regulators, for moving gates and dampers, etc. The pneumatic piston executive mechanisms alluded to frequently have either constant, or slightly varying, loads along their entire path of motion.

1. Unloaded Piston Executive Mechanism

The schematic of a piston executive mechanism is given in Fig. 1. Supply pressure P_0 is kept constant, pressure $P_3 = 1$ atmosphere. It is easily seen that, in the case when there are no loads on the piston and the piston has zero mass, the piston executive mechanism is an ordinary pneumatic flow chamber. By bearing in mind that the design of valves and other forms of distributing devices always provides equality of the geometric areas of the valve apertures, and making the additional simplifying assumption that in such mechanisms the discharge coefficients for the apertures for charging and discharging the cylinder voids equal each other, i.e., that $\mu_1 = \mu_2$, we obtain the equality of the effective areas: $\mu_1 f_1 = \mu_2 f_2 = f$.

From the general theory of pneumatic flow chambers [6] we know that if the effective areas of the valving apertures are varied in such fashion that their ratio always remains constant and if, moreover, pressures P_0 and P_3 are also unchanged, then the pressure in the intervalve chamber P_1 will also be constant.

Bearing in mind the remarks made above, we write the equations describing the motion of an unloaded piston executive mechanism. For writing the equations we shall use the approximate expressions for the air discharge through the valve apertures which differ only slightly in their accuracy from the well-known formulae of

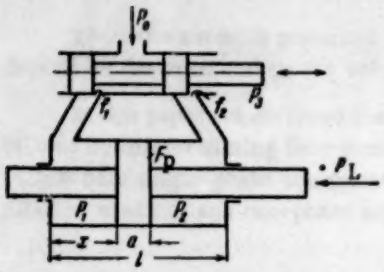

Fig. 1. Schematic of a pneumatic piston executive mechanism: P_0 is the supply pressure, P_3 is atmospheric pressure, f_1 and f_2 are the geometric areas of the apertures, F_p is the piston area, P_1 and P_2 are the pressures on the cylinder surfaces, P_L is the load on the piston rod, x is the piston coordinate, a is the piston thickness and l is the cylinder length.

Sen-Venana Vantsel' for an adiabatic process. The equations cited are obtained by equating the air discharge through the valve apertures to the air discharge along the cylinder cross section.*

Then, for the case of connected modes of flow through the valve apertures** "sub-sub," we shall have

for the intake chamber

$$f\sqrt{\frac{2g}{RT}P_1(P_0-P_1)}=\frac{P_1}{RT}F_p\dot{x}; \tag{1}$$

for the discharge chamber

$$f\sqrt{\frac{2g}{RT}(P_1-1)}=\frac{P_1}{RT}F_p\dot{x}, \tag{2}$$

where g is the acceleration due to gravity, R is the gas constant, T is the absolute temperature of the air ***, F is the piston area and x is the piston coordinate.

By eliminating pressure P_1 from Eqs. (1) and (2) and then substituting its value in one of them, we shall have:

$$\dot{x}=\frac{\sqrt{2gRT}}{F_p}\sqrt{\frac{2P_0}{P_0-1+\sqrt{(P_0-1)^2+4}}-1}\,f=k\varphi_1 f. \tag{3}$$

Here,

$$k=\frac{\sqrt{2gRT}}{F_p},\quad \varphi_1=\sqrt{\frac{2P_0}{P_0-1+\sqrt{(P_0-1)^2+4}}-1}.$$

In the case of combined modes of flow "sub-super," we have

$$f\sqrt{\frac{2g}{RT}P_1(P_0-P_1)}=\frac{P_1}{RT}F_p\dot{x}, \tag{4}$$

$$fP_1\sqrt{\frac{g}{2RT}}=\frac{P_1}{RT}F_p\dot{x}. \tag{5}$$

By carrying out an analogous transformation, we get

$$\dot{x}=\frac{k}{2}f. \tag{6}$$

It follows from (3) and (6) that the velocity of the piston executive mechanism for subcritical flow through the first and second valve apertures of the distributing device depends on the area of the valve aperture cross section, on the supply pressure and on the coefficient k and, in the case of subcritical flow through the first valve aperture and supercritical flow through the second, depends only on the effective area f of the valve apertures and on the coefficient k, and does not depend on the supply pressure P_0.

We now determine the areas in which Eqs. (3) and (6) are applicable.

*We shall assume that, prior to the moment under consideration, some motion of the piston had occurred, and that the transient response connected with establishing the pressure P_1, caused by the given pressure P_0, had terminated, i.e., we consider the steady-state motion.

**Here "sub-sub" and further "sub-super" denote, respectively, subcritical flow through the first and second valve apertures, and subcritical flow through the first, and supercritical flow through the second.

***We assume that the transition from the chambers before the valve apertures to the chambers after the valve apertures is isothermal, i.e., that the air temperatures in these chambers are all equal.

If we solve Eqs. (5) and (4) jointly, we find the following relationship, which is constant and equal to $P_1/P_0 = 0.8$ for subcritical flow through the first valve aperture and supercritical flow through the second.

We now investigate the function φ_1 which enters into expression (3).

By differentiating this function with respect to P_0 and equating the result to zero, we determine the pressure P_0 for which the function assumes its maximum value. This pressure will equal 2.5 atmospheres, and the value of function φ_1 for this pressure will be 0.5. Thus, with $P_0 = 2.5$ atmospheres, expression (3) turns out to be identical to expression (6) for the case "sub-super." As the source pressure is increased above the value of 2.5 atmospheres, the magnitude of function φ_1 will be decreased although, naturally, the combined modes of flow through the valve apertures will be "sub-super."

Therefore, for all values of P_0 greater than 2.5 atmospheres, expression (6) should be used or, what amounts to the same thing, instead of φ_1, the value 0.5 should be substituted in (3). The combined modes of flow through the valve apertures "super-sub" and "super-super" can not exist. This will be shown below for the more general case of a piston executive mechanism with a constant load.

2. Executive Mechanism with a Constant Load on the Piston

We now obtain the equations for a piston executive mechanism found under a constant load.

The weights of the air (Fig. 1) in the left and right chambers of the servocylinder will be, respectively,

$$Q_1 = xF_p \frac{P_1}{RT} \text{ and } Q_2 = (l - x - a) F_p \frac{P_2}{RT}.$$

By differentiating both equations with respect to time, we obtain the equations of the intake and discharge cylinder voids. By adding, in addition, a third equation which expresses the equality of the forces on the piston, we will have a system of equations which completely describes the law of motion of the piston rod:

$$\frac{dQ_1}{dt} = \frac{F_p}{RT} (\dot{x}P_1 + x\dot{P}_1), \tag{7}$$

$$\frac{dQ_2}{dt} = \frac{F_p}{RT} [-\dot{x}P_2 + (l - x - a)\ \dot{P}_2], \tag{8}$$

$$P_1 - P_2 = \frac{P_L}{F_p} = \Delta P. \tag{9}$$

Here, the derivatives $dQ_1/dt = f\psi(P_1, P_0)$ and $dQ_2/dt = f\psi(P_2, P_3)$ are the mass per second discharges of air through the valve apertures of the distributing device.

By analogy with the unloaded piston executive mechanism, and by also taking into account that the difference P_1-P_2 is a constant quantity, we shall assume that the pressures P_1 and P_2 are themselves constant.* In other words, of all the possible solutions to the system of (7), (8) and (9), we shall choose such for which P_1 and P_2 are constant.

Taking into account the assumptions made, we rewrite Eqs. (7), (8) and (9), for example, for the case "sub-sub," in the following form:

$$f\sqrt{\frac{2g}{RT} P_1 (P_0 - P_1)} = \frac{F_p}{RT} P_1 \dot{x}, \tag{10}$$

$$f\sqrt{\frac{2g}{RT} (P_2 - 1)} = \frac{F_p}{RT} \dot{x}P_2, \tag{11}$$

$$P_1 - P_2 = \Delta P \tag{12}$$

*Here, as previously, it is assumed that prior to the moment of consideration of the system's behavior there was a certain displacement of the piston under the action of a given supply pressure, and that the transient response connected with the establishment of the steady-state pressures P_1 and P_2 is terminated.

or

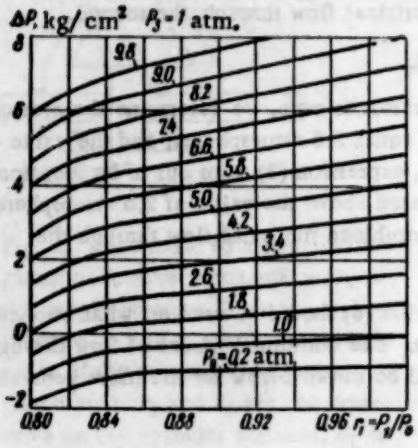

Fig. 2.

$$\dot{x} = k\sqrt{\frac{P_0}{P_1} - 1}\, f, \tag{13}$$

$$\dot{x} = k\frac{\sqrt{P_1 - \Delta P - 1}}{P_1 - \Delta P} f. \tag{14}$$

It might be possible here, by eliminating pressures P_1 and P_2 from the system of (10), (11) and (12), to find the equations of motion of the piston executive mechanism. However, such a route is a complicated one. It is more reasonable to use a graphicoanalytic method. This method reduces to the following. From the system of (10), (11) and (12), one determines ΔP as a function of P_1 and the supply pressure P_0. After this, one constructs the graphs of this function and, for given values of P_0 and ΔP, one finds P_1, substitution of which in (13) or (14) can give an expression for the velocity of piston motion.

We now find the dependence of ΔP on P_0 and P_1. For this, we shall replace the pressures, for greater generality, by their ratios, namely,

$$r = \frac{P_3}{P_0} = \frac{1}{P_0}, \quad r_1 = \frac{P_1}{P_0}, \quad r_2 = \frac{P_2}{P_1}, \quad r_3 = \frac{P_3}{P_2} = \frac{1}{P_2}.$$

Then, in accordance with (10), (11) and (12), and adding (16), we shall have

$$\frac{1}{r_2} = \Delta P r_3 + 1, \tag{15}$$

$$r = r_1 r_2 r_3, \tag{16}$$

$$\frac{1 - r_1}{1 - r_2} = r_1 r_3. \tag{17}$$

If we now eliminate r_2 and r_3 and then solve the resulting quadratic equation for ΔP, we get

$$\Delta P = \frac{r_1[2(1 - r_1) - r]}{2r(1 - r_1)} \pm \frac{\sqrt{r_1(5r_1 - 4)}}{2(1 - r_1)}. \tag{18}$$

By considering expression (18) we can easily convince ourselves that the minus sign before the second term is meaningless. Indeed, if expression (18) is analyzed with the minus sign before the second term, one can see that, with increasing r_1 and, consequently, with increasing P_1, the drop across the piston ΔP is decreased, which contradicts the physics of the phenomenon. By equating the second term to zero we obtain the minimum value of r_1 for which the drop across the piston, ΔP, still has a real value. This magnitude will be $r_1 = 0.8$. By its definition, the greatest value of r_1 cannot be greater than unity. Thus, for combined regimes of flow "sub-sub," the magnitude of r_1 varies from 0.8 to 1.

For joint regimes of flow in the intake and discharge valves of "sub-super," the system of equations of motion of the piston executive mechanism will be:

$$\dot{x} = kf\sqrt{\frac{P_0}{P_1} - 1}, \tag{19}$$

$$\dot{x} = \frac{k}{2} f, \tag{20}$$

$$P_1 - P_2 = \Delta P. \tag{21}$$

It is clear from this system of equations that the velocity of the piston in this case depends neither on the load on the piston rod nor on the supply pressure.

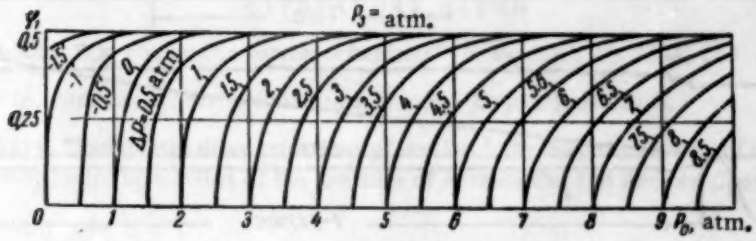

Fig. 3.

By solving (19) and (20) jointly, we find that $r_1 = 0.8$.

Thus, pressure P_1 and, consequently, also pressure P_2, does not depend on the velocity of the piston's motion. The joint modes of flow "super-sub" and "super-super" through the valve apertures of the cylinder's intake and outlet voids do not exist. It is possible to convince oneself of this by investigating, for the cases cited, the equations of motion of the piston executive mechanism. Then, for a joint regime of flow "super-sub" one obtains a complex drop on the piston and, for "super-super," $r_1 = 1$.

Since neither the one nor the other is possible, these modes of flow, with the assumptions made,* are absent in piston executive mechanisms.

Using Eq. (18) with respect to joint regimes of flow "sub-sub," and also taking into account the fact that, for the case "sub-super," $r_1 = 0.8$, we construct the graphs of Fig. 2.

All curves on the figure are for joint modes of flow "sub-sub," and all modes "sub-super" lie on the vertical line passing through the value of r_1 of 0.8.

By using these graphs we can, without previous knowledge of the joint modes of flow through the valve apertures in the intake and discharge chambers, but knowing only the supply pressure P_0 and the fall across the piston, $\Delta P = P_L/F_p$, determine the value of r_1. With this value of r_1 we can, either by formulae (13) or (14), compute the piston velocity.

To simplify even further the computations connected with the determination of the velocity of motion, x, of the piston, we can construct a graph (Fig. 3) whose axis of ordinates represents the function φ_1, equal to the radical in expressions (13) and (19), and whose axis of abscissas represents supply pressure P_0. The parameter is the pressure drop ΔP.

The graph was constructed by means of Fig. 2, from which the values of r_1 were determined, but the function φ_1 was computed from the formula $\varphi_1 = \sqrt{(1/r_1 - 1}$. All curves on the graph relate to the joint regimes of flow "sub-sub," and the horizontal line $\varphi_1 = 0.5$ relates to the joint regimes "sub-super." It is also clear from the graph that there will be no piston motion if $P_0 \leq (P_L/F_p + 1)$ atmosphere.

Knowing the value of φ_1, we can easily compute the velocity of the piston of the pneumatic piston executive mechanism from the formula

$$\dot{x} = k\varphi_1 f.$$

3. Experimental Verification of the Method for Determining the Velocity of Piston Rod Motion

To provide a basis for the correctness of the assumptions made, we experimentally determined the characteristics of a loaded piston executive mechanism. The piston executive mechanism investigated had a cylinder diameter of 31.6 mm and a total travel of 114 mm. The load was connected to its piston rod. With the motion due to gravity taken into account,** the load was so chosen that the pressure drop on the piston, ΔP, equalled

*Here we have kept in mind the equality of the effective areas of the valve apertures and the constancy of the load on the piston rod.

**For the piston executive mechanism investigated, the force of gravity was about 300 grams. Such a comparatively small gravitational force is explained by the use of rubber packing rings fitted on the piston with a significant initial gap.

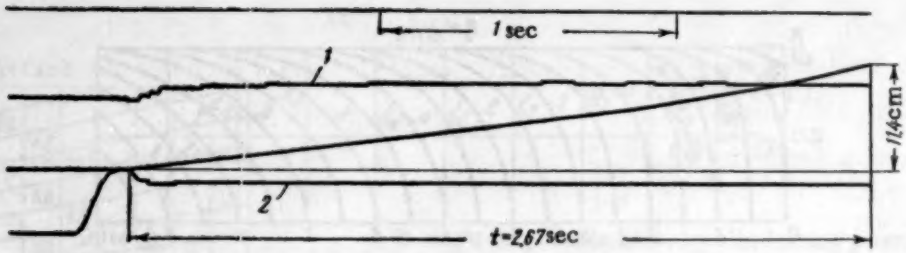

Fig. 4. Oscillogram showing the dependence of the pressures in the cylinder voids and of the piston travel on time. d is the diameter of the valve diaphragm, 1 is the pressure in the discharge chamber, 2 is the pressure in the intake chamber, P_0 = 2.1 atmospheres, d = 0.618 mm and x = 4.25 cm/sec.

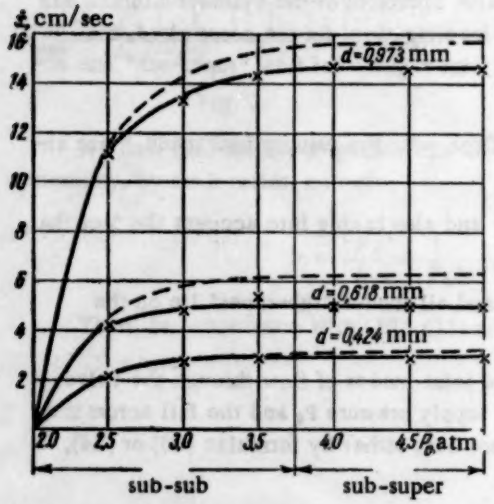

Fig. 5. Speed characteristics of the piston executive mechanism.

1 kilogram/cm^2. The piston rod load was implemented by a bob suspended by a fine steel wire drawn over light rollers. As valve elements we used impedances of the type "apertures in fine walls." In the pressure range investigated, the discharge coefficient of such valve elements is constant, and equal to 0.8 [8].

Two series of experiments were conducted. The first series of experiments was carried out in such fashion that the initial pressures in the cylinder voids equalled atmospheric pressure, and the supply pressure P_0 was applied to the chambers by jumps. However, in some case when there were large piston accelerations, caused by comparatively large valve diaphragm passage-way cross sections and large supply pressures, the pressure in the cylinder voids, with the lengths chosen for them, could not assume its steady-state value during the time of piston travel.

Since the methodology given in the paper is to be used only for computing steady-state piston motion, a second series of trials was made. The trials were carried out with an artificially established computed pressure in the cylinder voids.

The pressures, P_1 and P_2, in the cylinder voids, and the piston travel, were taken down on oscillograph paper.

In the course of the experiments, a significant number of oscillograms were collected for various supply pressures and different dimensions of the valve diaphragms. The results established that the pressures in the cylinder voids can be taken as constant, not depending on the dimensions of the valve apertures.

Moreover, the fact was also established that the supply pressure variations in the region "sub-super" do not, in practice, vary the magnitude of the piston velocity in the piston executive mechanism investigated.

As an illustration, Fig. 4 shown an oscillogram. The parameters with which the experiments were conducted are also shown there.

Figure 5 gives graphs constructed from the experimental results which express the dependence of piston rod velocity $\dot{x}$ of the piston executive mechanism on the supply pressure P_0 for various values of valve diaphragm diameters, with a pressure drop on the piston equal to 1 kg/cm^2. For purposes of comparison, the figure also shows, via the dotted curves, the computed curves.

As is obvious from Fig. 5, the maximum error does not exceed 7.6%*

*The maximum discrepancy between the computed and the experimental data (we are speaking now of the case when there were large valve diaphragm passage-way cross sections and large supply pressures) when the experiments were carried out without the establishment of the initial pressures in the cylinder voids was about 30%.

LITERATURE CITED

[1] V. N. Chelomei, "On pneumatic servomechanisms" [In Russian]. AN SSSR, Otd. Tekhn. Nauk 5 (1954).

[2] H. Levenstein, Pneumatic Servomechanisms, Control Engng. June, 1955.

[3] N. S. Gorskaya, "Pneumatic drive equations of motion " [In Russian]. Trudy of the First Scientific-Technical Conference of Junior Specialists of the Institute of Automation and Remote Control of the AN SSSR. Izd. AN SSSR (1957).

[4] L. L. Artobolevskii, E. V. Gerts, A. E. Kobrinskii and N. P. Raevskii, "Theoretical and Experimental investigation of a pneumatic drive for welding machines " [In Russian]. Trudy Seminar TMM 56. Izd. AN SSSR (1955).

[5] L. L. Artobolevskii, E. V. Gerts, A. E. Kobrinskii and N. P. Raevskii, "On the dynamics of pneumatic devices," [In Russian]. Trudy Seminar TMM 13, 51 (1953).

[6] L. A. Zalmanzon, "Graphs for determining the parameters of steady-state air flow through a system of calibrated apertures in pneumatic regulators," [In Russian]. Automation and Remote Control (USSR) 13, 2 (1952).

[7] G. T. Berezovets, V. N. Dmitriev and E. M. Nadzhafov, "On the admissible simplifications in the design of pneumatic regulators," [In Russian]. Priborostroenie 4 (1957).

[8] V. N. Dmitriev, "Computing the static characteristics of a pneumatic relay," [In Russian]. Automation and Remote Control (USSR) 17, 9 (1956).*

Received June 9, 1958

*See English translation.

ON ONE METHOD OF DESIGNING TWO-INPUT ELECTRONIC FUNCTION GENERATORS

V. B. Smolov

(Leningrad)

This paper deals with circuits for electronic function generators for simulating functions of two arguments, these circuits being constructed on the basis of standard operational amplifiers with controlled impedances in the input and feedback circuits.

Concrete block schematics are given for generators to realize two-argument functions which are characteristic of standard analog (simulation) technology.

1. Modern electronic simulating (analog) devices contain, as their basic standard elements, the so-called operational amplifiers, on the basis of which one can implement various mathematical operations on the input dc voltages. However, in simulating complicated functions of two variables [1-3]

$$U_{\text{out}} = \Phi(U_{\text{in}1},\ U_{\text{in}2}) \tag{1}$$

the electrical circuits of the corresponding electronic function generators, using standard operational amplifiers and constant impedances, are significantly complicated, leading to reductions in a number of the usage and technical indicators applied to analog devices.

In the present work we consider one of the possible methods of constructing function generators which are useful in simulating a number of important particular cases of function (1) with an accuracy, speed and reliability sufficient for practical use. The basic special feature of the generators to be considered is the presence in their circuits of standard elements from electronic digital technology by means of which the necessary switching of the operational amplifiers' input and feedback circuits is implemented.

2. The standard circuits for operational amplifiers [4, 5] provide two types of transformation of functional characteristics:

a) a standard operational amplifier with anode-grid negative feedback

$$z = -r_{20}\sum_{1}^{m}\frac{x_{2j}}{r_{2j}}, \tag{2}$$

b) an operational amplifier with differential inputs

$$z = r_{20}\sum_{0}^{m}\frac{1}{r_{2j}}\cdot\frac{\sum_{1}^{n}\frac{x_{1k}}{r_{1k}}}{\sum_{1}^{n}\frac{1}{r_{1k}}} - r_{20}\sum_{1}^{m}\frac{x_{2j}}{r_{2j}}. \tag{3}$$

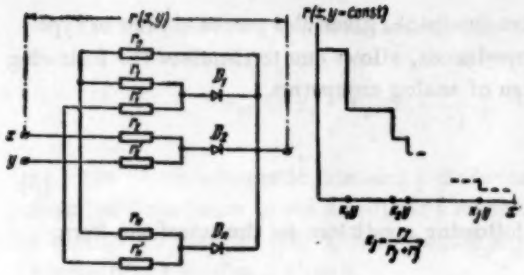

Fig. 1.

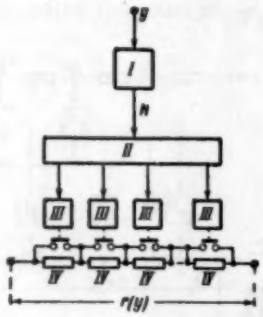

Fig. 2.

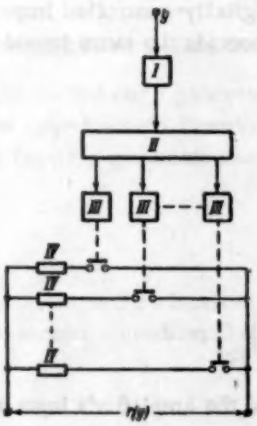

Fig. 3.

In formulae (2) and (3) we have used the notation: x_{1k} and x_{2j} are the input voltages of the amplifiers, r_{1k} and r_{2j} are the impedances in their input circuits and r_{20} is the impedance of the feedback circuit.

If the amplifier impedances are implemented in the form of variable ("controlled") ohmic impedances whose magnitude depends on the two input voltages

$$r_{1k} = r_{1k}(x_{1k},\ y_{1k}),\ r_{2j} = r_{2j}(x_{2j},\ y_{2j}), \qquad (4)$$

then the output voltages z of the amplifiers are complex functions of the input voltages; particular cases of these functions include the simulation of transformations of the type in (1) which are widespread in practice.

Thus, the use of controlled impedances with the characteristics of the form given in (4) in the block schematics of the operational amplifiers permits the solution of the problem of designing function generators with two inputs.

3. One can consider three fundamental circuits, constructed with controlled ohmic impedances, which are used in computing and measuring technology.

The first of these, shown in Fig. 1, [5] is used to obtain a functional impedance which varies stepwise in dependence on the input voltage x with constant values of voltage y:

$$r = r(x,\ y = \text{const}).$$

With a variable value of voltage y, the impedance retains its stepwise character, but is gradually shifted along the x axis.

The second fundamental circuit [6] for a controlled impedance contains linear transformer I which converts input voltage y into the digital quantity N which is statisticized in counter II. The output buses of the counter's flip-flops, one for each order (bit), excite the corresponding relays III whose contacts provide the successive switching of constant impedances IV (Fig. 2).

In the general case, the total magnitude of the digitally controlled impedance r can depend functionally on voltage y if, instead of a linear transformer of the input voltage to a digital quantity, one uses a nonlinear transformer.

The third circuit, shown in Fig. 3, is made up of constant ohmic impedances IV, connected in parallel in a common circuit by means of the contacts of the relays III which are switched by the output voltages of the order flip-flops of pulse counter II.

By combining the circuits considered with their impedances controlled by diodes (Fig. 1) and digitally (Figs. 2 and 3) with constant capacitive impedances in various block schematics of operational amplifier connections, one can provide simulation of various functions of type (1).

Below, we consider the block schematics of concrete modes of operation of nonlinear two-input transformers, constructed of the most widely used operational amplifiers with anode-grid negative feedback, and we provide certain information related to the questions of using block schematics with differential inputs.

4. Standard operational amplifiers with anode-grid negative feedback, given the proper choice of type and circuit for connecting the diode- and digitally-controlled impedances, allows one to simulate the following particular cases of function (1) which are widely used in the design of analog computers.

a) A multiplier

$$z = xy. \tag{5}$$

For the multiplication mode it is necessary to provide the following conditions for the transition from general relationship (2):

$$r_{20} \equiv y, \quad x_{21} \equiv x, \quad r_{21} = \text{const}, \quad (x_{2j})_{j>1} = 0, \quad \left(\frac{1}{r_{2j}}\right)_{j>1} = 0. \tag{6}$$

b) A divider

$$z = \frac{x}{y}. \tag{7}$$

The conditions for the transition from relationship (2):

$$r_{20} = \text{const}, \quad x_{21} \equiv x, \quad r_{21} \equiv y, \quad (x_{2j})_{j>1} = 0, \quad \left(\frac{1}{r_{2j}}\right)_{j>1} = 0. \tag{8}$$

Thus, the block schematic of the multiplier must contain one linear digitally-controlled impedance in the amplifier's feedback circuit, while the block schematic of the divider must contain the same impedance but in the amplifier's input circuit.

c) A multiplier-divider

$$z = \frac{x}{y} w. \tag{9}$$

For the transition from relationship (2):

$$r_{20} \equiv w, \quad r_{21} \equiv y, \quad x_{21} \equiv x, \quad (x_{2j})_{j>1} = 0, \quad \left(\frac{1}{r_{2j}}\right)_{j>1} = 0. \tag{10}$$

The block schematic contains linear digitally-controlled impedances in the amplifier's input and feedback circuits.

d) A multiplier for the function

$$z = yf(x). \tag{11}$$

For the transition from relationship (2):

$$r_{20} \equiv y, \quad r_{21} \equiv \frac{x}{f(x)}, \quad x_{21} \equiv x, \quad (x_{2j})_{j>1} = 0, \quad \left(\frac{1}{r_{2j}}\right)_{j>1} = 0. \tag{12}$$

The block schematic contains one diode-controlled impedance in the input circuit and one linear digitally-controlled impedance in the feedback circuit.

e) A multiplier for the function

$$z = f(x)\varphi(y). \tag{13}$$

For the transition from relationship (2):

$$x_{21} \equiv x, \quad r_{21} \equiv \frac{x}{f(x)}, \quad r_{20} \equiv \varphi(y), \quad (x_{2j})_{j>1} = 0, \quad \left(\frac{1}{r_{2j}}\right)_{j>2} = 0. \tag{14}$$

The block schematic contains a diode-controlled impedance in the input circuit and a functional digitally-controlled impedance in the amplifier's feedback circuit. The latter can be constructed in accordance with the block schematic of Fig. 3, or by combining diode-controlled and linear digitally-controlled impedances in the amplifier's feedback circuit.

a) If, in the general case, the amplifier's input circuit is formed of a group of diode-controlled impedances connected in parallel, $f_k(x_k)$, and the feedback circuit by a group of series-connected functional digitally-controlled impedances, $\varphi_j(y_j)$, then the block schematic of the function generator under consideration can be used to simulate a complicated multiplier-divider relationship on the sums of the function $f_k(x_k)$ and $\varphi_j(y_j)$.

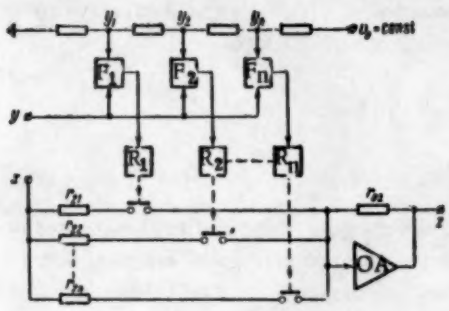

Fig. 4.

This latter circumstance allows one to use the generators for simulating functions of many variables by using the method of superposition.

5. If the function of two variables to be simulated

$$z = f(x, y),$$

does not belong to one of the previously considered particular cases, the following technical paths can be recommended for implementing it by means of digital-analog function generators based on operational amplifiers with controlled impedances.

a) One carries out a piece-wise linear approximation of the given function by one of its arguments and a piece-wise step-function approximation by the other argument. In correspondence with the method of approximation, the function generator must have, at its input, n diode-controlled impedances

$$r_{2j} \equiv \frac{x}{f(x, y_j)} \quad (j = 1, 2, 3, \ldots, n), \tag{15}$$

whose connections in the circuit are made in dependence on the ordinal number, j, of the function, i.e., on the value of the second argument, since

$$j = \Psi(y).$$

The function generator's block schematic is shown in Fig. 4.

b) A piece-wise linear approximation of the given function is made with respect to each of its arguments, for which the current value of the function is determined by the formula

$$z \approx f(x, y_j) + [f(x, y_{j+1}) - f(x, y_j)] \frac{y - y_j}{y_{j+1} - y_j}, \tag{16}$$

where $f(x, y_j)$ and $f(x, y_{j+1})$ are the values of the function z for the ends of the intervals of decomposition by argument y and $r_{2j} = f(x, y_j)$, y_{j+1} and y_j are the values of the interval of decompisition by argument y, and x and y are the current values of the arguments.

Since the function generator which simulates the function $z = f(x, y)$ in correspondence with formula (16) must provide the operation of multiplication, it is advantageous to connect a linear digitally-controlled impedance for argument y in the amplifier's feedback circuit.

In order to simplify the generator circuit, the decomposition should be uniform for y and nonuniform for x, so that the multiplication of the function increment Δf by the argument increment Δy within the limits of the

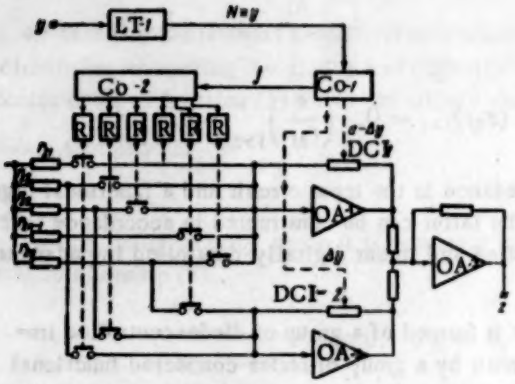

Fig. 5.

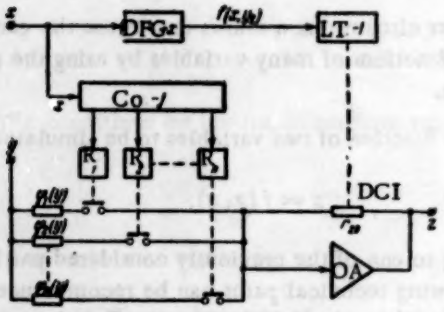

Fig. 6.

constant portion of the decomposition, $y_{j+1} - y_j$, is simply implemented technically by the block schematic described in section 4, a of this paper.

The over-all block schematic of the function generator which implements the simulation of functions by the method of piecewise approximations, and which realizes a variation of relationship (16)

$$z = f(x, y_j)\left(1 - \frac{\Delta y}{a}\right) + f(x, y_{j+1})\frac{\Delta y}{a}, \quad (17)$$

$$\Delta y = y - y_j, \quad a = y_{j+1} - y_j = \text{const},$$

is given in Fig. 5. The circuit contains three operational amplifiers, two of which (OA-1 and OA-2) serve to simulate relationships of the form

$$f_j\left(1 - \frac{\Delta y}{a}\right) \text{and} f_{j+1}\frac{\Delta y}{a},$$

while the third carries out the operation of summation.

The input voltage y is applied to linear transformer LT-1 which transforms the voltage to a digital quantity which is applied to pulse counter CO-1. The number of orders (places) in counter CO-1 corresponds to the constant interval of decomposition by argument y, so that counting of pulses occurs only in the interval a, after which the counter is cleared, and the overflow pulse, fixing the ordinal number j of the y decomposition, is applied to counter CO-2 of the segments a and then to the switching of the relay for the choice of the ordinal number of the function

$$f_j = f(x, y_j).$$

Multiplication of f_j by $a - \Delta y$, and of f_{j+1} by Δy, in the limits of the segment of the y decomposition is carried out by applying the inverted and the direct codes of the increment Δy to linear digitally-controlled impedances, respectively, DCI-1 and DCI-2.

c) The relationship to be simulated, $z = f(x, y)$, is presented in the form

$$z = f(x, y_1) + [f(x, y_2) - f(x, y_1)]\,\varphi_j(y). \quad (18)$$

If we choose

$$j = \Psi(x), \quad f(x, y_1) = 0, \quad (19)$$

then

$$z \approx f(x, y_2)\,\varphi_j(y). \quad (20)$$

The function generator which realizes (20) must contain, in the amplifier's input circuit, a group of diode-controlled impedances $\varphi_j(y)$, switched as a function of the magnitude of argument x and, in the feedback circuit, a functional digitally-controlled impedance, $r_{28} \equiv f$.

One of the possible block schematics for this function generator is shown in Fig. 6.

The functional digitally-controlled impedance is formed here by using a diode function generator, not containing an amplifier, in the x line.

d) In many cases, the given function z = f (x, y) can, with sufficient accuracy, be represented in the form of a sum

$$z = xf(y) + \Phi(x, y), \tag{21}$$

where

$$\Phi(x, y) \ll Z_{\max}. \tag{22}$$

In this case, the first term is simulated by the generator described in section 4, a and the second term, for example, by means of the block schematic of Fig. 6 which, by virtue of inequality (22), contains an insignificant number of elements.

6. For constructing a complicated integrating-differentiating two-input function generator, it is necessary to have in the amplifier circuits, in addition to controlled ohmic impedances, constant or controlled capacitive impedances.

For example, to simulate a function of the form

$$z = K_1 x + K_2 \frac{xw}{y} + K_3 w \frac{dx}{dt} + \\ + K_4 \int \frac{x}{y} dt \tag{23}$$

the following transition from relationship (2) is carried out:

$$x_{21} \equiv x, \quad r_{21} \equiv y, \quad r_{22} = \frac{1}{pC_1} = \text{const}, \quad r_{20} = Aw + \frac{1}{pC_2}, \\ (x_{2j})_{j>1} = 0, \quad \left(\frac{1}{r_{2j}}\right)_{j>1} = 0. \tag{24}$$

The block schematic of the function generator must contain linear digitally-controlled impedance r_{21} in the input circuit, shunted by capacitance C_1 and, in the feedback circuit, a series-connected linear digitally-controlled impedance Aw plus capacitance C_2.

For the construction of an integrating-differentiating function generator of higher order, it is necessary to use the block schematic of an amplifier with differential inputs; this same circuit is optimal for simulating functions of two variables which reduce to fractional-nonlinear relationships of the form

$$z = x \frac{f_1(x) f_2(y)}{[f_2(y) + f_3(x)][f_1(x) + f_4(y)]}. \tag{25}$$

As follows from (3), relationship (25) occurs for $(x_{2j})_{j>1} = 0$.

SUMMARY

The method considered for constructing electronic function generators with two inputs uses, as a rule, one operational amplifier and standard hardware of contemporary digital machines: linear coding transformers, counters, logical gates, which can be implemented with semiconducting and ferrite elements.

The accuracy and stability of these generators are determined basically by the accuracy and stability of two components, namely, the operational amplifier and the coding transformer.

The method possesses sufficient flexibility and universality from the point of view of the class of functions to be simulated and the accuracy of simulation.

The function generator circuits considered are significantly simplified if one argument is given in the form of a voltage, and the other in the form of a numerical equivalent.

The method permits the linking of the components of digital and analog electronic computers without

intermediate coding and decoding transformers, which can lead to a simplification of the construction of hybrid machines. Therefore, the function generators might find application in the design of controlling and measuring systems, and in instruments of various purposes.

The method can only be used for simulating nonlinear integral-differential functions if one of the arguments is given by a numerical equivalent.

LITERATURE CITED

[1] G. Korn and T. Korn, Electronic Analog Computers. [Russian translation of the English original], IL (1954).

[2] Meissinger, Convent Rec. IRE, pt. 4 (1955).

[3] A. L. Manukhin, A. A. Fel'dbaum and L. N. Fitsner, "Nonlinear two-input function generators," [In Russian]. PTE 4 (1957).

[4] G. L. Takhvanov, "Universal summing amplifiers," [In Russian]. Priborostroenie 4 (1957).

[5] A. D. Talantsev, "Designing diode function generators," [In Russian]. Automation and Remote Control (USSR) 17, 2 (1956).*

[6] V. F. Maiorov, Electronic Regulators. [In Russian]. GITTL (1956).

Received February 19, 1959

*See English translation.

ON THE DETERMINATION OF CERTAIN ERRORS OF ANALOG COMPUTERS

N. L. Sosenskii

(Moscow)

The paper considers the effect of the nonlinearity of the frequency characteristics of electronic analog computers' (EAC) operational elements in the solution of linear differential equations with constant coefficients. To determine the errors in solution we use a graphicoanalytic method based on the use of the logarithmic frequency characteristics. An example is given of the determination of the error in solving a second-order equation for the case when the operational amplifier has a frequency characteristic of complicated form and possesses a number of parasitic elements.

One of the important sources of errors in solving differential equations on an EAC is the nonlinearity of the frequency characteristics of the EAC elements. This error shows up most strikingly in solving differential equations with a high degree of conservatism (Cf., for example, [1]).

Works [1-3] are devoted to the error engendered by the frequency characteristics of the EAC elements in solving linear differential equations with constant coefficients. Determination of the error by the methods described in [1-3], for operational elements based on multistage operational amplifiers with arbitrary frequency characteristics and a host of parasitic elements (capacitance between the summing points and ground, output impedances, etc) leads to such horrendous computations that, in the cases considered, the use of these methods for computing the error turns out to be impossible in practice.

In this paper, for the determination of the frequency errors of an EAC, we employ the method, widely used in automatic control theory, of logarithmic amplitude (LAC) and logarithmic phase (LPC) characteristics. This allows one, by using the corresponding graphicoanalytic computations, to estimate the errors in cases where the operational amplifier's frequency characteristic is complicated, and also to take its parasitic elements into account.

Determination of EAC Errors Caused by its Operational Elements' Frequency Characteristics

We consider a linear differential equation with constant coefficients

$$y^{(n)} + \sum_{i=1}^{n} a_i y^{(n-i)} = f(t). \tag{1}$$

The block schematic for the model corresponding to Eq. (1) is shown in Fig. 1. The use in this circuit of operational elements with nonideal frequency characteristics leads to the result that the model solves, not the given Eq. (1), but an equation of higher order, and with other coefficients. We now estimate the difference of the natural motions corresponding to this equation from the natural motions corresponding to the original Eq. (1).

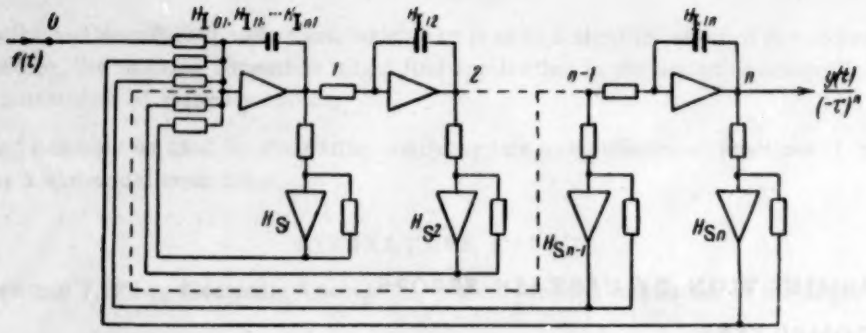

Fig. 1. Block schematic of a model for solving an n'th order linear differential equation with constant coefficients.

Initially, let the circuit of Fig. 1 consist of ideal elements with transfer functions

$$H_{Ik} = H_I = -\frac{1}{p\tau}, \quad H_{Si} = (-1)^{i+1} a_i \tau^i, \tag{2}$$

where H_{Ik} is the transfer function from the k'th input of the i'th integrator to its output, H_{Si} is the transfer function of the i'th scale amplifier, p is the complex frequency and τ is the integrators' time constant.

A plus sign in a transfer function H_{Si} indicates that the corresponding scale amplifier must be replaced by a resistance. We shall consider the circuit of Fig. 1 as a system with a multichannel feedback at whose input (at point 0) is applied the stimulus f(t). The system's output is the point at which the solution of the equation is obtained (point n). We shall denote the transfer functions of the portions of the circuit of Fig. 1 from point 0 to points 1, 2,..., n by Y_1, Y_2,..., Y_n, and we shall assume that all the poles of these functions lie in the left half of the complex p plane.

Now, let the circuit of Fig. 1 consist of nonideal elements (Fig. 2). Then, the elements' transfer functions are defined by expressions (III)-(XIII) (cf. Appendix I). The logarithmic frequency characteristics of the transfer functions of the nonideal elements can be easily constructed from the transfer function, A (p), of the operational amplifier (which is considered as being given) by means of nomograms for closed-loop systems [4]. We denote the transfer functions of the portion of the circuit with nonideal elements from point 0 to points 1, 2,..., n by $\overline{Y}_1$, $\overline{Y}_2$,..., $\overline{Y}_n$, and we assume that all the poles of these functions lie in the left half-plane, and also that the LAC and the LPC of the functions $\overline{Y}_i$, for the frequencies $\omega < \omega^*$, differ little from the LAC and LPC of the functions Y_i. Here, ω^* is the boundary frequency which satisfies the conditions, $\omega^* > 1/T_i$, $\omega^* > \omega_i$, and $1/T_i$ and $-\eta_i\omega_i \pm j\omega_i\sqrt{1-\eta_i^2}$ are the poles of function Y_i. These assumptions are equivalent to the requirement that the error in solving the equation on the EAC be small.

The difference between $\overline{Y}_n$ (p) and Y_n (p) may amount to a shift of the already existing poles, or to the appearance of new poles or zeroes. We denote by p_i the poles of function Y_n (p), and by $\overline{p}_i$ and $\overline{p}_k$, respectively, the shifted former poles and the new poles of function $\overline{Y}_n$ (p) (i = 1, 2,..., n; k = n + 1, n + 2,..., n + m). In the case of Eq. (1), to each of the poles p_i there corresponds, in the natural motions, a term of the form $D_i e^{p_i t} + \varphi_i$. For the natural motions of the model, there corresponds to each of the poles $\overline{p}_i$ and $\overline{p}_k$ a term of the form $\overline{D}_i e^{\overline{p}_i t} + \overline{\varphi}_i$ or $\overline{D}_k e^{\overline{p}_k t} + \overline{\varphi}_k$. With this, the quantities D_i and φ_i are uniquely defined by the initial conditions and by the poles p_i, and the quantities $\overline{D}_i$, $\overline{\varphi}_i$, $\overline{D}_k$ and $\overline{\varphi}_k$ by the initial conditions, by the poles $\overline{p}_i$ and $\overline{p}_k$ and by the new functions $\overline{Y}_n$. In the present work we do not consider the errors arising from the discrepancies between $\overline{D}_i$ and D_i, or between $\overline{\varphi}_i$ and φ_i. Thus, to determine the error in solution, it suffices to know the values of the new poles and the magnitude of the displacement of the old poles. This information can be obtained by a comparison of the logarithmic characteristics corresponding to the function Y_i and $\overline{Y}_i$, which are constructed by successively applying the nomograms for closed systems to the circuit of Fig. 1. For this, in correspondence with the assumptions made above,

$$|\overline{H}(\omega)| = |H(\omega)|\,[1 + \delta(\omega)], \quad \arg\overline{H}(\omega) = \arg H(\omega) + \alpha(\omega), \tag{3}$$

$$|\overline{Y}_i(\omega)| = |Y_i(\omega)|\,[1 + \varepsilon_i(\omega)], \quad \arg\overline{Y}_i(\omega) = \arg Y_i(\omega) + \gamma_i(\omega), \tag{4}$$

where δ, α, ε_i, γ_i are small deviations from the ideal characteristics.

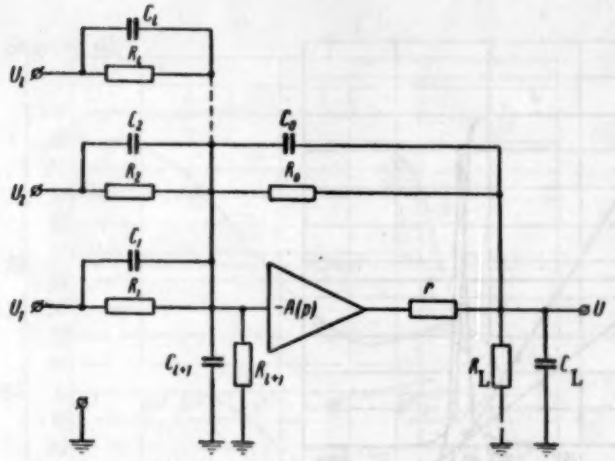

Fig. 2. Equivalent circuit of the EAC's operational elements.

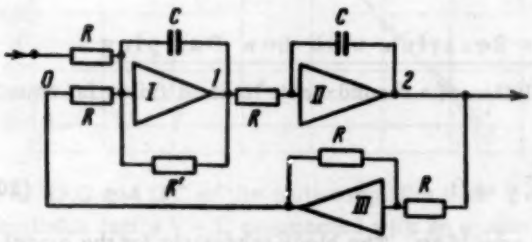

Fig. 3. Block schematic of the model for solving a second-order equation.

It may turn out that certain of the quantities δ, α or ϵ_i and γ_i are so small that it is difficult to determine them from the corresponding curves. Relationships are given in Appendix II which permit the determination of the quantities $\delta(\omega)$ and $\alpha(\omega)$ from the frequency characteristics of the operational amplifier for the frequency range in which $|G(j\omega)/H(j\omega)| \gg 1$.

The functions $G(j\omega)$ and $H(j\omega)$ are defined by expressions (IV), (VIII), (X) and (XIII) of Appendix I. Also found there are the expressions for determining $\epsilon_i(\omega)$ and $\gamma_i(\omega)$ from the frequency characteristics of the functions $\overline{Y}_{i-1}(j\omega)$ and $Y_{i-1}(j\omega)$ for the frequency range in which $|\overline{Y}_{i-1}\overline{H}_i\overline{H}_{Si}| \gg 1$.

We now consider how, from a comparison of the frequency characteristics $\overline{Y}_n$ and Y_n, we can determine the new poles, and also the translations of the old poles. The appearance of new poles is expressed by the fact that, at frequencies greater than boundary frequency ω^*, the LAC of function $\overline{Y}_n$ will have a slope larger than 20 n db/decade and may also have a maximum (the LAC of function Y_n, at frequencies greater than ω^* has a slope of 20 n db/decade and has no maxima). It is easily shown that all the new real, $-1/T_k$, and complex, $-\eta_k\omega_k \pm j\omega_k\sqrt{1-\eta_k^2}$, poles of function $\overline{Y}_n$ are characterized by the relationships

$$\frac{1}{T_k} > \Omega,\quad \omega_k = \overline{\omega}_{km},\quad \eta_k = \frac{1}{2\overline{Y}_{km}}, \tag{5}$$

where Ω is the angular frequency at which the slope of the LAC changes, $\overline{\omega}_{km}$ is a frequency at which a new maximum of the LAC is observed and $\overline{Y}_{km}$ is the value of this maximum.

It is also easily shown that a shift of a real pole $-1/T_i$ leads, in the LPC, to a variations with maximum value, for $\omega \approx 1/T_i$, equal to

$$d\varphi_m = \frac{dT_i}{2T_i}. \tag{6}$$

This same shift, at frequencies $\omega > 1/T_i$, leads to a variation in the amplitude characteristic:

$$\frac{d|Y|}{|Y|} = -\frac{dT_i}{T_i}. \tag{7}$$

The shift of a pair of complex conjugate poles, $-\eta_i\omega_i \pm j\omega_i\sqrt{1-\eta_i^2}$, leads to a variation in the magnitude of Y_{im} and in the position, ω_{im}, of the corresponding maximum in the LAC. For $\eta_i^2 \ll 1$, these variations are defined by the formulae

$$\frac{d|Y_{im}|}{|Y_{im}|} \approx -\frac{d\eta_i}{\eta_i}, \tag{8}$$

$$\frac{d\omega_{im}}{\omega_{im}} \approx \frac{d\omega_i}{\omega_i}. \tag{9}$$

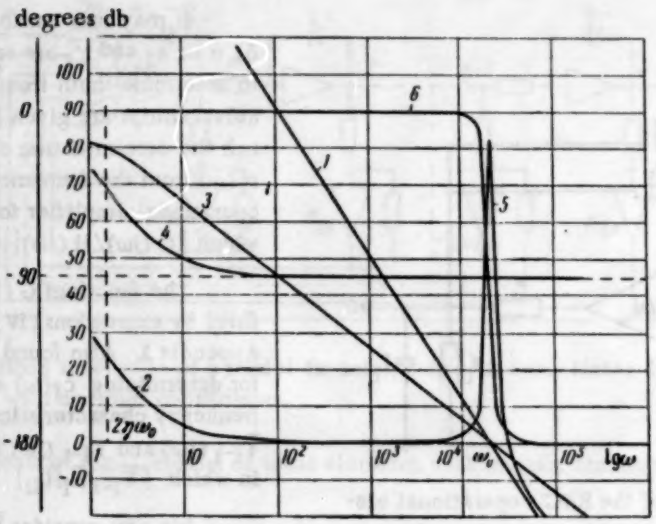

Fig. 4. LAC and LPC of the model consisting of ideal elements.

Determination of the Error for a Second-Order Equation with Low Damping

As an example, we now consider the error in the EAC solution of a second-order linear differential equation with low damping and constant coefficients,

$$y'' + 2\eta\omega_0 y' + \omega_0^2 y = 0 \tag{10}$$

in the case when the operational amplifiers are three-stage dc amplifiers. The block schematic for the model corresponding to Eq. (10) is given in Fig. 3 ($RC = 1/\omega_0$, $R'C = 1/2\eta\ \omega_0$).

One easily convinces oneself that

$$Y_1 = -\frac{1}{2\eta}\frac{2\eta\omega_0}{p + 2\eta\omega_0}, \tag{11}$$

$$Y_2 = -\frac{Y_1\omega_0}{p + Y_1\omega_0} = -\frac{\omega_0^2}{p^2 + 2\eta\omega_0 p + \omega_0^2}, \tag{12}$$

$$|Y_{2m}| = \frac{1}{2\eta}. \tag{13}$$

Figure 4 gives the frequency characteristics corresponding to the block schematic with ideal elements: 1) $|\Phi|$, 2) arg Φ, 3) $|Y_1|$, 4) arg Y_1, 5) $|Y_2|$ and 6) arg Y_2. The LAC and LPC of the function $Y_2(j\omega)$ were constructed by means of the nomograms for closed systems from the LAC and LPC of the function $\Phi(j\omega) = \omega_0 Y_1(j\omega)/j\omega$.

Let the LAC and LPC of the operational amplifiers used in the block schematic of Fig. 3 be given by graphs 1 and 2 of Fig. 5. These are standard curves for a three-stage dc amplifier with stabilizing circuits. If the parasitic parameters are given then, by using the nomograms for closed systems and the formulae of Appendix I, we can successively construct the LAC and LPC for the transfer functions of the scale amplifier and the integrators, and also for the functions $\bar{Y}_1$ and $\bar{Y}_2$. Figure 5 shows the frequencies of the scale amplifier:

$$1 - |A| - 95\ \text{db},\ 2 - \arg A,\ 3 - \left|\frac{A}{1 + 2\cdot 10^{-9} j\omega}\right| - 115\ \text{db},$$

$$4 - \arg\frac{A}{1 + 2\cdot 10^{-9} j\omega},\quad 5 - \left|1 - \frac{1 + 2\cdot 10^{-9} j\omega}{10A}\right|,\quad 6 - \arg\left[1 - \frac{1 + 2\cdot 10^{-9} j\omega}{10A}\right],$$

$$7 - \left|\frac{G_S}{H_S}\right| - 88\ \text{db},\ 8 - \arg\frac{G_S}{H_S},\ 9 - |\bar{H}_S|,\ 10 - \arg\bar{H}_S.$$

Figure 6 gives the frequency characteristics of the integrator:

$$1 - |A| - 95\ \text{db},\quad 2 - \arg A,\quad 3 - \left|1 - \frac{1 + 5j\omega}{10^6 A}\right|,\quad 4 - \arg\left[1 - \frac{1 + 5j\omega}{10^6 A}\right]$$

$$5 - \left|\frac{G_I}{H_I}\right| - 94\ \text{db},\quad 6 - \arg\frac{G_I}{H_I},\quad 7 - |H_I|,\quad 8 - \arg\bar{H}_I,\quad 9 - |\bar{Y}_1|,$$

$$10 - \arg\bar{Y}_1.$$

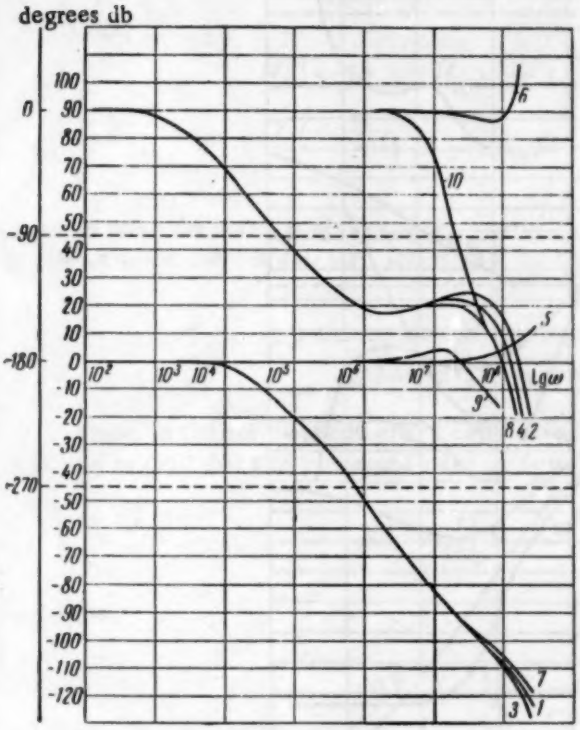

Fig. 5. LAC and LPC of the scale amplifier with transmission factor k = 1, constructed with an operational amplifier with transfer function A (jω).

Figure 7 gives the frequency characteristics corresponding to the block schematic with nonideal elements: 1) $|\bar{Y}_1\bar{H}_I\bar{H}_S|$, 2) arg $\bar{Y}_1\bar{H}_I\bar{H}_S$, 3) $|\bar{Y}_2|$, 4) arg $\bar{Y}_2$.

All curves constructed are for the following values of the basic parameters of the circuit of Fig. 3 and of the parasitic elements of the integrators and the scale amplifier (the parasitic elements have been taken as being identical which, obviously, does not destroy the generality of the discussion).

1. The angular frequency of the natural motions corresponding to Eq. (10): $\omega_0 = 2 \cdot 10^4$/sec, the damping decrement is $\eta = 3.6 \cdot 10^{-5}$.

2. The parameters of the elements of the circuit in Fig. 3: $R = 10^3$ ohm, $R' = 1.4 \cdot 10^7$ ohm, $C = 5 \cdot 10^{-8}$ farad.

3. The leakage impedance on the integrator condensers is $R_0 = 10^8$ ohm, the output impedance of the operational amplifiers is $r = 10^5$ ohm, the load impedance for all amplifiers is $R_L = R = 10^3$ ohm, the leakage impedance with the summing points and ground is $R_2 = 10^5$ ohm.

4. The parasitic capacities are: between the input and the summing point, $C_1 = 2$ nanofarad between the output and the summing point of the scale amplifier, $C_0 = 2$ nanofarad, between the summing point and ground, $C_2 = 5$ nanofarad; the load capacitance on the operational element is $C_L = 20$ nanofarad.

In correspondence with formulae (XXI)-(XXIV), the small deviations from the ideal characteristics [cf. (3) and (4)] will equal

$$\delta_S(\omega) = -\left|\frac{H_S}{G_S}\right|\cos\varphi_S, \quad \alpha_S(\omega) = \left|\frac{H_S}{G_S}\right|\sin\varphi_S + \omega R(C_1 - C_0), \tag{14}$$

$$\delta_I(\omega) = -\left|\frac{H_I}{G}\right|\cos\varphi_I, \quad \alpha_I(\omega) = \left|\frac{H_I}{G_I}\right|\sin\varphi_I + \omega RC_1 + \frac{1}{\omega R_0 C}, \tag{15}$$

$$\varepsilon_1(\omega) = \delta_I(\omega), \quad \gamma_1(\omega) = \alpha_I(\omega), \tag{16}$$

where $\varphi_S(\omega) = \arg G_S/H_S$ and $\varphi_I(\omega) = \arg G_I/H_I$.

For a system with very low damping, determination of the maximum of the amplitude characteristic by means of the nomograms for closed systems is difficult, since the nomograms are not usually constructed for very small values of the phase characteristic of the open-loop system's transfer function. Therefore, to determine the magnitude of $|\bar{Y}_{2m}|$, we use the obvious expression

$$\bar{Y}_2 = -\frac{1}{\bar{H}_S}\,\frac{\bar{Y}_1\bar{H}_I\bar{H}_S}{\bar{Y}_1\bar{H}_I\bar{H}_S - 1}. \tag{17}$$

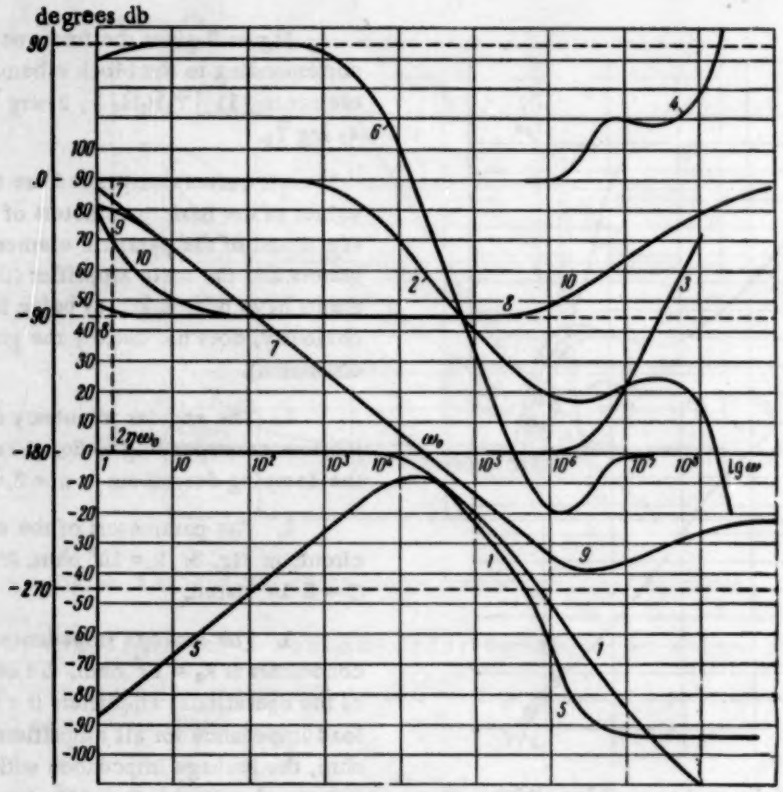

Fig. 6. LAC and LPC of the integrator with time constant $\tau = 1/\omega_0$, constructed with an operational amplifier with transfer function A (jω).

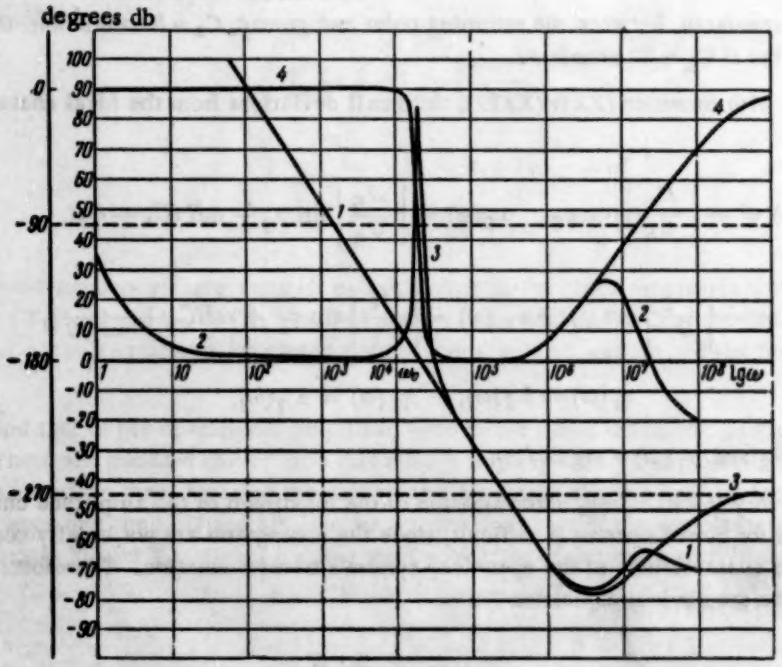

Fig. 7. LAC and LPC of the model consisting of nonideal elements.

Here,

$$|H_S| = 1, \quad \arg H_S = -\pi, \quad |H_I| = \frac{\omega_0}{\omega}, \quad \arg H_I = -\frac{3\pi}{2},$$

$$|Y_1| = \frac{\omega_c}{\omega}\frac{\omega}{\sqrt{\omega^2 + (2\eta\,\omega_0)^2}}, \quad \arg Y_1 = -\frac{3\pi}{2} + \operatorname{arc\,tg}\frac{2\eta\omega_0}{\omega}.$$

By using the notation of (3) and (4), expanding the individual terms of expressions (17) in series and ignoring terms of the order of $|H/G|^2$ and η^2 (all the quantities α, γ, δ_I, ϵ_I have order $|H/G|$), we obtain

$$\bar{Y}_2 = \frac{1-\delta_S}{\left[1-\left(\frac{\omega}{\omega_0}\right)^2(1-\delta_S-\delta_I-\varepsilon i)\right] + j\left[\left(\frac{\omega}{\omega_0}\right)^2\left(\alpha_I+\gamma_1+\frac{2\eta\,\omega_0}{\omega}\right)+\alpha_S\right]}. \tag{18}$$

Since, in the neighborhood of $\omega = \omega_{2m}$, the quantity $|G/H| \gg 1$, i.e., all the quantities α, γ, δ_I, $\epsilon_I \ll 1$, it is easy to show that the maximum value of Y_2 occurs for

$$\bar{\omega}_{2m} = \omega_0\left[1 - \frac{\delta_S(\omega_{2m}) + \delta_I(\omega_{2m}) + \varepsilon_1(\omega_{2m})}{2}\right] \tag{19}$$

and equals

$$|\bar{Y}_{2m}| = \frac{1}{\alpha_S(\omega_{2m}) + \alpha_I(\omega_{2m}) + \gamma_1(\omega_{2m}) + 2\eta}. \tag{20}$$

In the case under consideration we obtain, from the corresponding characteristics,

$$\delta_S(\omega_{2m}) = -\frac{\cos(-60°)}{1.3\cdot10^4} = -4\cdot10^{-5}, \quad \alpha_S(\omega_{2m}) = \frac{\sin(-60°)}{1.3\cdot10^4} = -6.9\cdot10^{-5},$$

$$\delta_I(\omega_{2m}) = -\frac{\cos(-20°)}{1.8\cdot10^4} = -5.3\cdot10^{-5},$$

$$\alpha_I(\omega_{2m}) = \frac{\sin(-20°)}{1.8\cdot10^4} + 4\cdot10^{-5} + \frac{1}{10^5} = 3.1\cdot10^{-5},$$

$$\varepsilon_1(\omega_{2m}) = -5.3\cdot10^{-5}, \qquad \gamma_1(\omega_{2m}) = 3.1\cdot10^{-5}.$$

By using (19) and (20) we find that $\bar{\omega}_{2m} \approx 2\cdot10^4$/sec and $|\bar{Y}_{2m}| = 84$ db.

From a comparison of the characteristics of Fig. 7 with the characteristics of Fig. 4, we can draw the following conclusions.

1. Transfer function $\bar{Y}_2$ has a pair of complex conjugate poles which are translated from their position in Y_2, leading to a variation of the damping decrement and of the frequency of the natural motions. In correspondence with (13), the damping decrement, instead of being $\eta = 3.6\cdot10^{-5}$, equals $\bar{\eta} = 3.2\cdot10^{-5}$. In correspondence with (19), the frequency equals $\bar{\omega}_0 \approx 2\cdot10^4$ sec.

2. Transfer function $\bar{Y}_2$ has no new poles, as compared to Y_2.

3. By the appropriate choice of the operational amplifiers' frequency characteristics one can compensate for the EAC's frequency errors with an accuracy determined by the constancy of the operational amplifier's gain at zero frequency. The same result can be attained by using the well-known methods for correcting the phase characteristics, or by choosing a sufficiently large operational amplifier gain at the operating frequency.

An error analysis for the case of very small values of operational amplifier output impedance (< 1 ohm), carried out by the method described above, showed that the transfer function $\bar{Y}_2$ in this case, in addition to a shifted pair of complex poles, had new poles, defined by expressions (5). The additional terms in the natural motions which correspond to the new poles have the form $e^{-\eta_k\omega_k t}\cos\omega_k t$, and are damped by a factor of e during time $t = 1/\eta_k\omega_k = 2\bar{Y}_{km}/\bar{\omega}_{km}$, where $\bar{\omega}_{km}$ is the frequency approximately equal to the operational amplifier's cut-off frequency (i.e., the frequency for which $|A(j\omega)| = 1$) and $\bar{Y}_{km}$ is a quantity defined by the

stability margin of the operational amplifiers. Thus, a larger velocity of damping of the additional terms in the natural motions can be provided by a sufficiently high cut-off frequency and by a larger stability margin of the operational amplifiers.

SUMMARY

A graphico-analytic method which is based on the use of logarithmic frequency characteristics allows one to estimate, from given graphs of the operational amplifiers' LAC and LPC, the error, due to the frequency properties of the model's elements, in solving linear differential equations with constant coefficients. With this, one can determine graphically the fraction of the total error attributable to each source of error in the solution, and can also determine the possible measures to take for decreasing and compensating the errors.

The author wishes to express his gratitude to A. A. Kuz'min and A. A. Vasil'ev for their discussion of a number of questions which arose during preparation of this paper.

APPENDIX I

Expressions for the Transfer Functions of the Integrator and the Summing Amplifier with Account Taken of the Parasitic Elements

Figure 2 gave the equivalent circuit for the operational element with account taken of the parasitic capacities, amplifier output impedance and condenser leakage impedance. By using the method of nodal voltages for the amplifier adding points and the output points, we get

$$U = -\frac{\sum_{i=1}^{l} \frac{R_0}{R_i}\frac{1+pR_iC_i}{1+pR_0C_0}U_i}{1-\frac{R_0r(g_{in}+pC_{in})(g_{out}+pC_{out})}{\left[-A(p)+\frac{r}{R_0}+prC_0\right](1+pR_0C_0)}}, \tag{I}$$

where

$$g_{in}=\sum_{i=1}^{l+1}\frac{1}{R_i},\quad g_{out}=\frac{1}{r}+\frac{1}{R_0}+\frac{1}{R_L},\quad C_{in}=\sum_{i=0}^{l+1}C_i,\quad C_{out}=C_0+C_L. \tag{II}$$

The transfer function from the summer's k'th input to its output can be written in the form

$$\overline{H}_{kS}(p)=H_{kS}\frac{1+p\tau_k}{1+p\tau_0}\frac{1}{1+\frac{H_{kS}}{G_{kS}}}, \tag{III}$$

where

$$G_{kS}(p)=-G_{k0}\frac{A(p)(1+p\tau_0)}{(1+p\tau_{in\cdot S})(1+p\tau_{out\cdot S})}\left[1-\frac{r}{R_0}\frac{1+p\tau_0}{A(p)}\right], \tag{IV}$$

$$G_{k0}=\frac{1}{R_krg_{in}g_{out}}=\frac{1}{\left(1+\frac{r}{R_0}+\frac{r}{R_L}\right)\sum_{i=0}^{l+1}\frac{R_k}{R_i}}, \tag{V}$$

$$\tau_0=R_0C_0,\quad \tau_k=R_kC_k, \tag{VI}$$

$$\tau_{in\cdot S}=\frac{C_{in}}{g_{in}},\quad \tau_{out\cdot S}=\frac{C_{out}}{g_{out}}, \tag{VII}$$

$$H_{kS}=-\frac{R_0}{R_k}. \tag{VIII}$$

The transfer function of the integrator's k'th input to its output is written in the form

$$\overline{H}_{k\mathrm{I}}(p) = H_{k\mathrm{I}} \frac{1 + p\tau_k}{1 + \frac{1}{p\tau_0}} \frac{1}{1 + \frac{H_{k\mathrm{I}}}{G_{k\mathrm{I}}}}, \tag{IX}$$

where

$$G_{k\mathrm{I}}(p) = -G_{k0} \frac{A(p)\left(1 + \frac{1}{p\tau_0}\right)}{(1 + p\tau_{\mathrm{in}\cdot\mathrm{I}})(1 + p\tau_{\mathrm{out}\cdot\mathrm{I}})} \left[1 - prC_0 \frac{1 + \frac{1}{p\tau_0}}{A(p)}\right], \tag{X}$$

$$\tau_{\mathrm{in}\cdot\mathrm{I}} = \frac{C_{\mathrm{in}}}{g_{\mathrm{in}}} = R_k C_0 \frac{1 + \sum_{i=1}^{l+1} \frac{C_i}{C_0}}{\sum_{i=0}^{l+1} \frac{R_k}{R_i}}, \tag{XI}$$

$$\tau_{\mathrm{out}\cdot\mathrm{I}} = \frac{C_{\mathrm{out}}}{g_{\mathrm{out}}} = rC_0 \frac{1 + \frac{C_{\mathrm{L}}}{C_0}}{1 + \frac{r}{R_0} + \frac{r}{R_{\mathrm{L}}}}, \tag{XII}$$

$$H_{k\mathrm{I}} = -\frac{1}{pR_kC_0}. \tag{XIII}$$

APPENDIX II

Derivation of the Expressions for the Deviations of the Frequency Characteristics of the Nonideal Elements from those of the Ideal Elements

We consider the transfer function of some system with a single-channel feedback

$$\overline{H}(p) = H(p)\,h(p) \frac{1}{1 + \frac{H(p)}{G(p)}} \tag{XIV}$$

in the frequency band where

$$\left|\frac{G(j\omega)}{H(j\omega)}\right| \gg 1.$$

We introduce the notation

$$G(j\omega) = |G|e^{j\varphi},\ H(j\omega) = |H|e^{j\psi},\ \overline{H}(j\omega) = |\overline{H}|e^{j\overline{\psi}},\ h(j\omega) = |h|e^{j\beta}. \tag{XV}$$

By substituting in (XIV) we get

$$|\overline{H}| = \frac{|Hh|}{\sqrt{\left[1 + \left|\frac{H}{G}\right|\cos(\psi - \varphi)\right]^2 + \left[\left|\frac{H}{G}\right|\sin(\psi - \varphi)\right]^2}}, \tag{XVI}$$

$$\overline{\psi} = \psi + \beta - \operatorname{arc\,tg} \left|\frac{H}{G}\right| \frac{\sin(\psi - \varphi)}{1 + \left|\frac{H}{G}\right|\cos(\psi - \varphi)}. \tag{XVII}$$

If we expand expressions (XVI) and (XVII) in powers of $|H/G|$, limiting ourselves to the first two terms of each series, we obtain

$$|\overline{H}| = |Hh|\left[1 - \left|\frac{H}{G}\right|\cos(\psi - \varphi) + B_1\right]. \tag{XVIII}$$

$$\overline{\psi} = \psi + \beta - \left|\frac{H}{G}\right|\sin(\psi - \varphi) + B_2, \tag{XIX}$$

where B_1 and B_2 are the remaining terms of the series where, for $|G/H| \geq 10$,

$$B_1 \leq 2.5\left|\frac{H}{G}\right|^2, \quad B_2 \leq 1.7\left|\frac{H}{G}\right|^2. \tag{XX}$$

We neglect B_1 and B_2 and employ these expressions in the transfer functions of the EAC's operational elements, (III) and (IX). Then, by taking (VIII) into account and by ignoring terms of the order of $(\omega\tau_k)^2$ and $(\omega\tau_0)^2$, we obtain for the summer and for the scale amplifier

$$|\overline{H}_{kS}| \approx \frac{R_0}{R_k}\left[1 - \frac{R_0}{R_k|G_{kS}|}\cos(\varphi_{kS} - \psi_S)\right], \tag{XXI}$$

$$\overline{\psi}_{kS} \approx -\pi + \omega\tau_k - \omega\tau_0 + \frac{R_0}{R_k|G_{kS}|}\sin(\varphi_{kS} - \psi_{kS}). \tag{XXII}$$

For the integrator, by taking (XIII) into account and by neglecting terms of the order of $(\omega\tau_k)^2$ and $1/(\omega\tau_0)^2$, we get

$$|\overline{H}_{kI}| \approx \frac{1}{\omega R_k C_0}\left[1 - \frac{1}{\omega R_k C_0|G_{kI}|}\cos(\varphi_{kI} - \psi_I)\right]. \tag{XXIII}$$

$$\overline{\psi}_{kI} \approx -\frac{3}{2}\pi + \omega\tau_k + \frac{1}{\omega\tau_0} + \frac{1}{\omega R_k C_0|G_{kI}|}\sin(\varphi_{kI} - \psi_I). \tag{XXIV}$$

LITERATURE CITED

[1] A. B. MacNee, Some Limitations on the Accuracy of Electronic Differential Analyzers. Proc. IRE, vol. 40, No. 3 (1952).

[2] V. A. Marsocci, An Error Analysis of Electronic Analog Computers. Trans. IRE, vol. EC-5, No. 4 (1956).

[3] P. C. Dow, An Analysis of Certain Errors in Electronic Differential Analyzers: I) Bandwidth Limitations. Trans. IRE, vol. EC-6, No. 4 (1957).

[4] N. T. Kuzovkov, Automatic Control Theory [In Russian], Oborongiz (1957).

Received May 28, 1959

CHOOSING FLEXIBLE FEEDBACK PARAMETERS FOR MAGNETIC AMPLIFIERS BY MEANS OF INTEGRAL ESTIMATES

L. V. Safris
(Rostov-on-the-Don)

By means of the method of integral quadratic estimates, we elucidate the effect of the parameters of a flexible (velocity) feedback on the duration of the transient response in a choke-coupled magnetic amplifier.

INTRODUCTION

As is well known, to accelerate the transient responses in magnetic amplifiers, use is made of flexible feedback, implemented as a rule by means of an additional winding W_1 which is connected in parallel with the load via condenser C and impedance R_a (Fig. 1).

The recommendations to be found in the literature [1, 2] on the choice of parameters for the accelerating contour lead, principally, to the determination, in each individual case, of the value of C for which there would occur a critical aperiodic of damping oscillatory transient response. With such a posing of the problem there remains hidden the effectiveness of the flexible feedback, in the sense of its effect on the speed with which the transient response occurs. Elucidating this effectiveness by constructing curves of the increase in load current i_L is a very arduous operation.

We mention that, in some cases, the attainment of a critical aperiodic mode may have little effect as far as accelerating the transient response is concerned.

The method, described in this paper, of choosing the parameters of the feedback path by means of an integral quadratic estimate of the form

$$J = \int_0^\infty (i_L - I_L)^2 dt \qquad (1)$$

(where I_L is the steady-state value of load current) possesses the advantage that it allows one simultaneously to determine the desirable acceleration of the transient response. As is well known [3], integral estimates are widely used for choosing the parameters of automatic control systems.

We shall consider as optimal those amplifier parameters for which the integral in (1) is a minimum.

For the basis of our investigation we take the method of linearizing the magnetization characteristic, described in [4]. From this same work we also carry over the following basic notation and relationships: W_θ is the number of turns in the control winding; W_ω is the number of turns in the ac winding; W_β is the number of turns in the current feedback winding; $\beta = W_\beta / W_\theta$; Q is the cross section of the two cores; l is the mean length of the magnetic circuit; R_θ is the impedance of the control circuit; R_L is the load impedance; r_a is the forward impedance of a rectifier bridge arm; $R = R_L + r_a$; i_θ (I_θ) are control currents; i_L (I_L) are load currents; i_a is the current in the accelerating feedback circuit; u_θ (U_θ) are the voltages applied to the input of the controlling

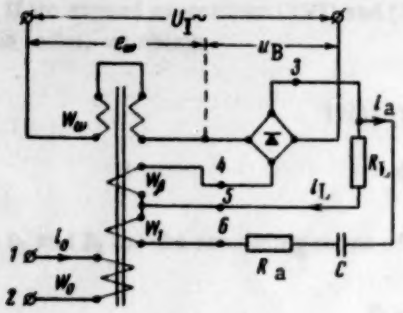

Fig. 1.

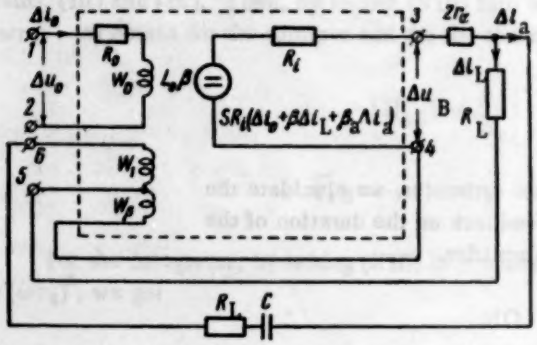

Fig. 2.

winding; e_ω (E_ω) are the "average" values of emf in the ac winding; u_B (U_B) are the average values of voltage on the rectifier bridge input terminals; B_0 is the constant (monotonic) component of magnetic induction; H_0 is the constant component of the magnetic field strength; μ_i, σ, μ_c and ν are coefficients which characterize the core properties and which are defined by the magnetization characteristics; $S = \sigma W_0/W_\omega$ is the amplifier characteristic's steepness; $R_i = \mu_i W_\omega^2 Q\omega/l$ is the amplifier's internal impedance; $k_i = R_i S/(R_i + R)$ is the current gain; $k_{i\beta} = k_i/(1 - \beta k_i)$ is the current gain with inflexible feedback present; k_p is the power gain with no feedback present; $L_{\omega\beta}$ is the amplifier's input inductance with an inflexible feedback present; Δ is a symbol which indicates that one has chosen the increment of the corresponding quantity within the limits of the characteristic's linear portion.

1. Equivalent Circuit and Basic Assumptions

To compute the integral quadratic estimate, we make use of the equivalent amplifier circuit shown in Fig. 2. The values of the equivalent amplifier circuit parameters were stated above in the list of nomenclature. The equivalent circuit is obtained on the basis of the conclusions presented in work [4], in which the following basic assumptions are made:

a) the values of the magnetic induction and the magnetic field strength do not go beyond the limits of the rectilinear region of the characteristic for simultaneous magnetization by constant and variable fields;

b) the load is active;

c) the impedance of the control circuit is sufficiently small:

$$\left(\frac{\beta^2 R_0}{R_i + R} \ll 1\right);$$

d) the rectifiers are replaced by forward constant impedances, and the back impedances are taken to be infinitely large;

e) the transient response occurs "slowly" in comparison with the supply voltage period;

f) with magnetization of the cores, the increments of voltage at the rectifiers' input terminals and of the emf in the ac winding are equal in magnitude and opposite in sign:

$$\Delta E_\omega = -\Delta U_B, \tag{2}$$

i.e., the load curve is approximated by straight line segments.

By keeping in mind the active load and the low impedance of the ac winding, we can agree that the first five assumptions carry over to the case under consideration here. As for the last assumption, it requires special verification since, in contradistinction to the cases treated in [4], the load is connected in parallel with a contour containing a capacitance.

Experimental investigations (cf. Appendix 1) allow the following points to be established.

In the majority of cases, for the steady-state mode, Eq. (2) remains valid with an adequate degree of accuracy even when there is capacitance present. As examples, one can take curves 1 (C = 0) and 2 (C = 4.7 microfarad) on Fig. 3 (cf. Appendix 2).

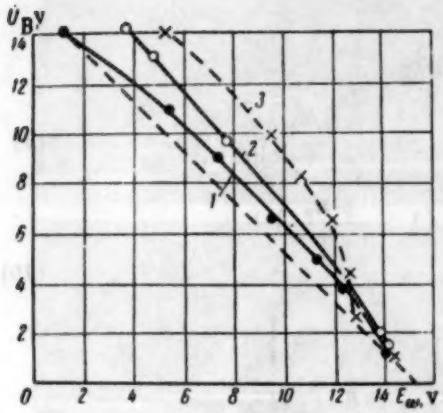

Fig. 3.

With certain values of capacitance, and with low forward rectifier impedance, the character of the function $E_\omega = f_1(U_B)$ changes somewhat. This case is illustrated by curve 3 on Fig. 3. Here, on individual segments, ΔU_B is significantly larger than the corresponding ΔE_ω. The reasons for such a character of the curves are considered in [5, 6].

In what follows, in those cases when, in a definite range of variation of U_B, equality (2) is significantly violated, we shall replace it by the equality

$$\Delta E_\omega \approx -\alpha \Delta U_B, \tag{2a}$$

where α will be determined experimentally.

The assumption of Eq. (2) is the basis for considering the equivalent circuit of Fig. 2 as being valid in the steady state.

We shall assume that expression (2) can be extended to the transient response, and we use the circuit of Fig. 2 for investigating the steady-state processes while remembering that the final results obtained with such assumptions have agreed well with experimental data.

As was established in [4], the increment of the constant component of the magnetic induction is defined by the expression

$$\Delta B_0 = \mu_c \Delta H_0 - \frac{\nu}{\omega} \frac{\Delta e_\omega}{W_\omega Q}. \tag{3}$$

For the equivalent circuit of Fig. 2, with Eq. (2) being taken into account, Eq. (3) takes the form:

$$\Delta B_0 = \frac{\mu_a W_0}{l}(\Delta i_0 + \beta \Delta i_L + \beta_a \Delta i_a) + \frac{\nu}{\omega} \frac{R_L \Delta i_L + 2r_a(\Delta i_L + \Delta i_a) + W_\beta Q \frac{d(\Delta B_0)}{dt}}{W_\omega Q}. \tag{4}$$

To determine the load current, Δi_L, we use the following equations, which follow from the equivalent circuit:

$$S(\Delta i_0 + \beta \Delta i_L + \beta_a \Delta i_a) = \Delta i_L + \Delta i_a + \frac{R_L \Delta i_L + 2r_a(\Delta i_L + \Delta i_a) + W_\beta Q \frac{d(\Delta B_0)}{dt}}{R_1}, \tag{5}$$

$$W_1 Q \frac{d(\Delta B_0)}{dt} + \frac{1}{C}\int \Delta i_a dt + R_a \Delta i_a = R_L \Delta i_L, \tag{6}$$

$$W_0 Q \frac{d(\Delta B_0)}{dt} + R_0 \Delta i_0 = u_0. \tag{7}$$

In (4) and (5), $\beta_a = (W_\beta + W_1)/W_0 = W_a/W_0$.

We note in passing that, by discarding the right-most term in Eqs. (4) and (5) and by replacing μ_c by $\mu_\Delta = \Delta B_0/\Delta H_0$, we obtain equations analogous to those which were taken in [1, 2] as the basis of the assumptions regarding the constancy of the input inductance and the validity of equating the variable and constant components of the ampere-turns in the transient response.

If we apply the Laplace transform to Eqs. (4)-(7) with zero initial conditions and a stepwise varying voltage u_0, and then solve these equations jointly, we find that

$$\Delta i_L(p) = k_{i\beta} \frac{U_0}{R_0} \frac{1}{p}\left[\frac{b_2}{a_2} + \frac{\left(b_1 - b_2 \frac{a_1}{a_2}\right)p + \left(1 - \frac{b_2}{a_2}\right)}{a_2 p^2 + a_1 p + 1}\right], \tag{8}$$

Here,

$$b_1 = R_g C - \frac{\beta R_0}{k_{i\beta}}\left(\frac{\tau_{v\beta}}{R} + \frac{\tau_{a\beta}}{R_i + R}\right),$$

$$b_2 = \frac{R_0 C}{k_{i\beta}}\left\{\tau_{v\beta}\left[n(\beta_c - \beta) - \frac{\beta R_a}{R}\right] + \tau_{a\beta}\left[m(\beta_a - \beta) - \frac{\beta R_a}{R_i + R}\right]\right\}, \quad (9)$$

$$a_1 = \tau_{v\beta}\left[1 - \beta R_0\left(\frac{1}{R k_{i\beta}} - \frac{1}{(R_i + R)k_i}\right)\right] + \tau_{a\beta}\left(1 + \frac{\beta^2 R_0}{R_i + R}\right) + C[R_c - (\beta_c k_{i\beta} - m_\beta)R_L], \quad (10)$$

$$a_2 = C\left\{\tau_{v\beta}\left[\frac{R_0}{k_{i\beta}}\left\{(\beta_a k_{i\beta} - m_\beta)\left[\beta_a - \beta\left(1 - \frac{R_L}{R}\right)\right] + n(\beta_a - \beta)\right\} + R_a\left[1 + \beta R_0\left(\frac{1}{(R_i + R)k_i} - \frac{1}{R k_{i\beta}}\right)\right] + nR_L\left(1 + \frac{\beta R_0}{(R_i + R)k_i}\right)\right] + \tau_a\left[R_0(\beta_a - \beta)[(\beta_a k_{i\beta} - m_\beta)\beta - \beta_a](1 - \beta k_i) + R_a\left(1 + \frac{\beta^2 R_0}{R_i + R}\right) + R_L\left(m + \frac{\beta R_0}{R_i + R}\right)\right]\right\}. \quad (11)$$

For brevity in writing Eqs. (9), (10) and (11), we introduced the following notation:

$$n = \frac{2r_a}{R}, \quad m = \frac{R_i + 2r_a}{R_i + R}, \quad m_\beta = \frac{m}{1 - \beta k_i}, \quad \tau_a = \frac{W_0^2 Q \mu_c}{l R_0},$$

$$\tau_{a\beta} = \frac{\tau_a}{1 - \beta k_i}, \quad \tau_v = \frac{\nu}{\omega}\frac{W_0}{W_\omega}\frac{R}{R_0}k_i, \quad \tau_{v\beta} = \frac{\tau_v}{1 - \beta k_i}.$$

The quantities $\tau_{a\beta}$ and $\tau_{v\beta}$ are the components of the control circuit's time constant ($\tau_{a\beta} + \tau_{v\beta} = L_{e\beta}/R_e$) which, respectively, do not, and do, depend on the load impedance.

To simplify the analysis, we assume that, with small values of R_e, the gain does not depend on the magnitude of capacitance C. Actually, as was noted in [2] the presence of the flexible feedback engenders a certain increase in k_i due to the even harmonics.

For ideal cores and diodes, when $\tau_a = 0$, $m = 1$, $n = 0$ and $R_i \gg R$, the coefficients of Eq. (8) take the form:

$$b_1 = R_g C - \frac{\beta R_0}{k_{i\beta} R_L}\tau_{v\beta}, \quad b_2 = -\frac{\beta R_0}{k_{i\beta} R_L} R_a C \tau_{v\beta}, \quad (9a)$$

$$a_1 = \tau_{v\beta}\left(1 - \frac{\beta R_0}{k_{i\beta} R_L}\right) - C[(\beta_a k_{i\beta} - m_\beta)R_L - R_a], \quad (10a)$$

$$a_2 = C\tau_{v\beta}\left[\frac{\beta_a R_0}{k_{i\beta}}(\beta_a k_{i\beta} - m_\beta) + R_a\left(1 - \frac{\beta R_0}{k_{i\beta} R_L}\right)\right]. \quad (11a)$$

2. The Expression for J and its Analysis

To compute the integral J, corresponding to the amplifier's stable mode, we use the method presented in [7].

We shall be interested only in the transitional component of the load current, so that the term b_2/a_2 in (8) is dropped. As a result, we obtain the expression

$$\Delta i_L^*(p) = k_{i\beta}\frac{U_0}{R_0}\frac{1}{p}\frac{b_{1p}^* + b_0^*}{a_2 p^2 + a_1 p + 1}, \quad (12)$$

where $b_1^* = b_1 - b_2 a_1/a_2$ and $b_0^* = 1 - b_2/a_2$.

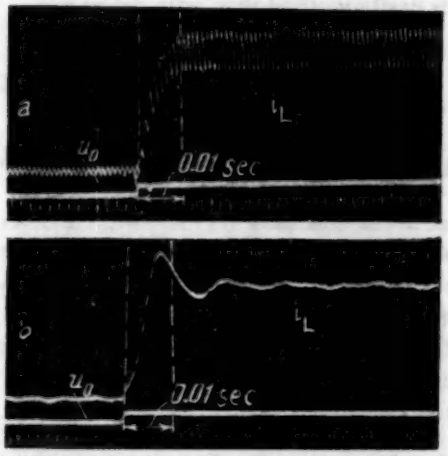

Fig. 4.

By retaining (with added primes) the notation of [7], we find

$$\Delta' = a_1 a_2 - 1,$$
$$\Delta_1' = \Delta_2, \quad \Delta_0' = a_1^2 a_2 + a_2, \quad B_0' = b_0^*, \quad B_1' = b_1^{*2}.$$

Consequently,]

$$J = \left(k_{i\beta}\frac{U_0}{R_0}\right)^2 \frac{a^2(a_1-b_1)^2 + (a_2-b_2)^2}{2a_1 a_2}. \tag{13}$$

We note that if C = 0 then $b_2 = 0$ and $a_2 = 0$. As a rule, $\beta R_0/(R_i + R)k_i \ll 1$ and $\beta^2 R_0/(R_i + R) \ll 1$. We can therefore assume that $a_1 - b_1 \approx a_1 \approx \tau_{v\beta} + \tau_{a\beta}$. Then,

$$J_{C=0} = \left(k_{i\beta}\frac{U_0}{R_0}\right)^2 \frac{\tau_{v\beta} + \tau_{a\beta}}{2}.$$

For subsequent analysis, it is convenient to consider the dimensionless quantity

$$\theta = \frac{J}{J_{C=0}} = \frac{a_2(a_1-b_1)^2 + (a_2-b_2)^2}{a_1 a_2(\tau_{v\beta} + \tau_{a\beta})}. \tag{14}$$

The coefficient θ allows one to judge how the duration of the transient response after introduction of the accelerating feedback changes from its value when no such feedback is present.

Expressions (8)-(11) and (14) and all their consequences remain valid when there is no rigid feedback (of current), if it is borne in mind that $\beta = 0$ and, consequently, $\tau_{v\beta}$, $\tau_{a\beta}$, $k_{i\beta}$ and m_β must be replaced by, respectively, τ_v, τ_a, k_i and $\underline{m}$.

If the numerical values of the amplifier parameters are substituted in (14), one can then determine the quantities C, β_a and R_a which correspond to the least value of θ and for which, consequently, one can expect the best flow of the transient response.

We begin our analysis with the particular, and simplest, case of an ideal amplifier. Since, in any amplifier, the power is $\beta R_0/k_{i\beta}R_L = \beta k_{i\beta} R_0/k_{i\beta}^2 R_L \ll 1$, it is then easily established from (9a) and (11a) that $a_2 - b_2 \approx a_2$. Thus, for the case under discussion,

$$\theta = \frac{(a_1-b_1)^2 + a_2}{a_1\tau_{v\beta}}. \tag{14a}$$

By substituting the values of b_1, a_1 and a_2 from (9a), (10a) and (11a), we can write expression (14a) in the following way:

$$\theta = \frac{[\tau_{v\beta} - c(\beta_a k_{i\beta} - m_\beta)R_L]^2 + C\tau_{v\beta}\left[\frac{R_0\beta_a}{k_{i\beta}}(\beta_a k_{i\beta} - m_\beta) + R_a\right]}{\tau_{v\beta}\{\tau_{v\beta} - C[(\beta_a k_{i\beta} - m_\beta)R_L - R_a]\}}. \tag{15}$$

We first explain under what conditions the flexible feedback can provide an acceleration of the transient response, i.e., when the condition $\theta < 1$ will hold. If we take the difference between the numerator and the denominator of expression (15), the condition $\theta < 1$ is transformed to the inequality

$$CR_L(\beta_a k_{i\beta} - m_\beta)\left[(\beta_a k_{i\beta} - m_\beta)CR_L - \tau_{v\beta}\left(1 - \frac{\beta_a R_0}{k_{i\beta}R_L}\right)\right] < 0. \tag{16}$$

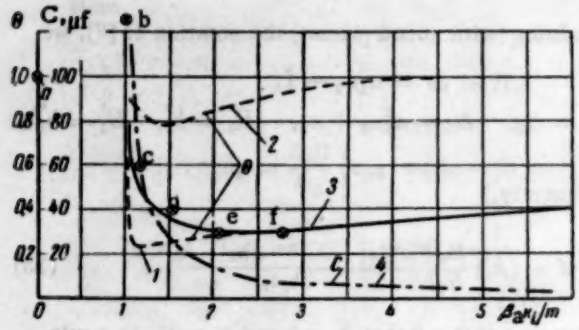

Fig. 5.

It is easily verified that, with the values

$$\beta_a k_{i\beta} - m_\beta = \frac{\beta_a k_i - m}{1 - \beta k_i} \leqslant 0 \qquad (17)$$

$\beta_a R_\theta / k_{i\beta} R_L \ll 1$ and inequality (16) does not hold. Since $m = 1$ and $k_1 \approx W_\theta / W_\omega$, condition (17) is equivalent to the condition $W_a / W_\omega \leq 1$. In other words, the effect of the flexible feedback will not be realized if the current branching off through the capacitor after mutliplication by a factor of $\beta_a k_1$ induces in the load a component which is less than Δi_a. With such a choice of parameters, the inclusion of the flexible feedback leads to a more of less delicate retardation of the transient response.

It is necessary to mention that the conditions for choosing parameters which were given in [2] are inadequate, since the appearance of a periodic component in the curve for current i_L does not always lead to an acceleration of the transient response.

As an example, Fig. 4 shows the oscillograms of the load current i_L, obtained in the mode when $\beta_a k_1 = m$ (cf. Appendix 3). As is obvious, the appearance of a periodic component (Fig. 4b) has no practical effect on the duration of the transient response, which remained the same as when there was no flexible feedback (Fig. 4a).

Numerical verification by means of expression (15) for a number of amplifiers, and also the experimentally-founded statements given in [2] allow one to assume that, all other things being equal, an increase in R_a leads to a retardation of the transient response. Therefore, it is always advantageous to have a minimal value of R_a (the resistance of winding W_1).

Thus, the problem of choosing the parameters of the flexible feedback reduces to determining the optimal, mutually corresponding, values of C and β_a.

In the general case, the answer can be obtained by solving the system of equations $\partial\theta/\partial C = 0$ and $\partial\theta/\partial\beta_a = 0$. In view of the complexity of the expressions for C and β we use the following route.

From the equation $\partial\theta/\partial C = 0$ we find the optimal value of capacitance as a function of β:

$$C_{opt} = \frac{\tau_{v\beta}}{(\beta_a k_{i\beta} - m_\beta) R_L - R_a}\left(1 - \sqrt{1 - \left[\frac{(\beta_a k_{i\beta} - m_\beta) R_L - R_a}{(\beta_a k_{i\beta} - m_\beta) R_L}\right]^2 + \frac{\tau_{v\beta}\left[\frac{\beta_a R_\theta}{k_{i\beta}}(\beta_a k_{i\beta} - m_\beta) + R_c\right][(\beta_a k_{i\beta} - m_\beta) R_L - R_a]}{\tau_{v\beta}[(\beta_a k_{i\beta} - m_\beta) R_L]^2}}\right). \qquad (18)$$

Due to the smallness of R_a in comparison with $(\beta_a k_{i\beta} - m_\beta) R_L$ (except in the neighborhood of points where $\beta_a k_{i\beta} = m_\beta$), the radicand can be simplified somewhat, and we take

$$C_{opt} = \frac{\tau_{v\beta}}{(\beta_a k_{i\beta} - m_\beta) R_L - R_a}\left(1 - \sqrt{\frac{\beta_a R_\theta}{k_{i\beta} R_L}}\right). \qquad (19)$$

By substituting (19) in (15), and taking certain simplifications into account, we find the relationship for $C = C_{opt}$

$$\theta(\beta_a) = 2\sqrt{\frac{\beta_a R_\theta}{k_{i\beta} R_L}} - \frac{\beta_a R_\theta}{k_{i\beta} R_L} + \frac{R_c}{(\beta_a k_{i\beta} - m_\beta) R_L - R_L}. \qquad (20)$$

Expression (20) allows one to compute the very best value of β_a, corresponding to the minimal value of θ.

As an example of such a construction, Fig. 5 shows, on curve 1 (cf. Appendix 4), the representation of

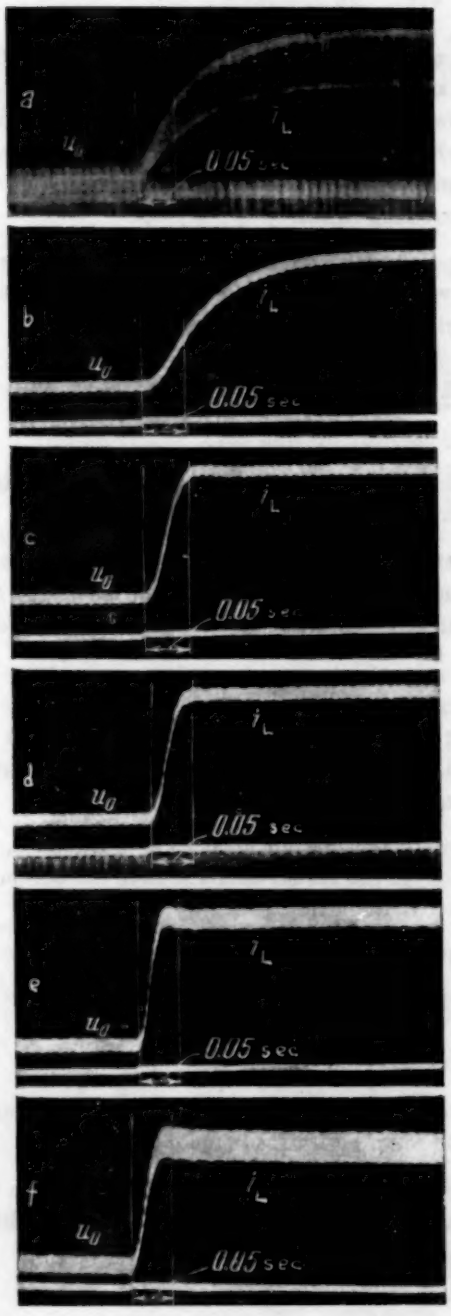

Fig. 6.

the function $\theta = f_1(\beta_a k_i / m)$ for $C = C_{opt}$. The same figure shows, by curve 2, $\theta = f_2(\beta_a k_i / m)$ for $C = C_{opt}$ for the same amplifier with no rigid feedback ($\beta = 0$).

A comparison of these curves established that the flexible feedback is significantly more effective when a rigid feedback is also present.

As is obvious from (20), for $R_a = 0$, and all other conditions being equal, the least duration of the transient response will be attained for

$$\theta = 2\sqrt{\frac{1-\beta k_i}{k_P}} - \frac{1-\beta k_i}{k_P}.$$

Here we have taken into account that the least value of β_a (for $C_{opt} \to \infty$) equals $1/k_i$. It follows from this expression that an increase in the rigid feedback coefficient β entails a decrease in θ, i.e., amplifies the accelerating action of the flexible feedback.

On the other hand, curves 1 and 2 of Fig. 5 attest to the fact that an extreme increase of β_a, even with the proper choice of capacitance, leads to an increase in θ.

We recall that we have been considering the case of an ideal core and an ideal diode.

We now explain how the value of the coefficient θ varies if we take the actual amplifier parameters into account.

To this end, we determine the expression for θ by means of (14) by substituting the numerical values of the amplifier parameters in (9), (10) and (11). Then, from the equation $\partial\theta/\partial C = 0$ we find C_{opt} as a function of β_a. Finally, by assigning different values to β_a, we compute $\theta = f_3(\beta_a k_i / m)$ for $C = C_{opt} = f_4(\beta_a k_i / m)$.

Curves 3 and 4 of Fig. 5 correspond to these last two functions (cf. Appendix 4).

As is obvious, curve 3 has a minimum which is significantly less sharply expressed than the one for curve 1, it being displaced to the region where $\beta_a k_i / m = 2$ to 3. Within the limits of these values of $\beta_a k_i / m$ there must occur a greatest velocity of transient response which will be less than for the ideal amplifier.

Curve 3 agrees well with experimental data. Figure 6 shows oscillograms a, b, c, d, e and f which illustrate the course of the transient response for different values of $\beta_a k_i / m$ when $C = C_{opt}$ (cf. Appendix 3). The values of capacitance were chosen experimentally. The parameters of the amplifier tested (cf. Appendix 1) were close to those for which curves 3 and 4 of Fig. 5 were calculated (cf. Appendix 4). The oscillograms correspond to points a, b, c, d, e and f on this figure

Thus, to the degree that the parameters of the amplifier under consideration deviate from those of the ideal amplifier, the effectiveness of the flexible feedback will be decreased.

If one wishes to avoid the difficult computations described above, one can take, as a first approximation, $\beta_a k_i = 2$ to 3 for high-quality cores, and $\beta_a k_i = 3$ to 4 for the rest. An approximate value of the corresponding C_{opt} can be found by means of the following approximate formula which follows from (18) for $\beta_a R_a / k_i \beta R_L \ll 1$:

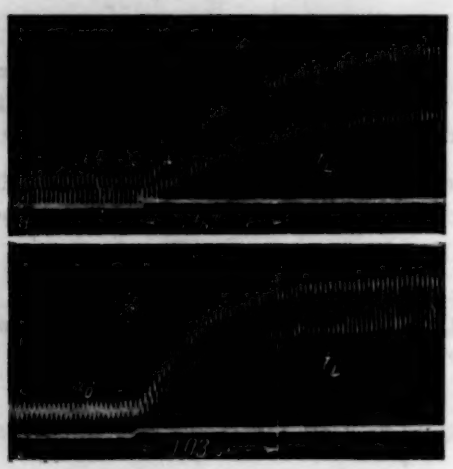

Fig. 7.

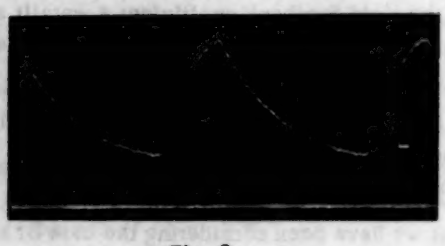

Fig. 8.

$$C_{\text{opt}} = \frac{\tau_v}{(\beta_a k_i - m) R_L}. \qquad (21)$$

In conclusion, we turn our attention to the fact that this same expression can be obtained if we start from the following considerations. We find a value of C such that the first derivative of the frequency characteristic, $i_L(j\Omega)$, in the neighborhood of the point $\Omega = 0$ would equal zero. Here, Ω denotes the frequency of the controlling signal.

By using expression (8) we find that $di_L(j\Omega)/d(j\Omega) = 0$ for $\Omega = 0$ if $a_1 - b_1 = 0$. By taking (9a) and (10a) into account, we find an expression for C which coincides with (21).

3. Peculiar Features of Certain Modes of Operation

The discussion given above remains qualitatively valid also for $\alpha < 1$ in formula (2a).

Quantitatively, naturally there are changes since the taking into account of expression (2a) leads to the change of a number of the coefficients of Eqs. (4) and (5), which depend on α (cf. Appendix 5).

In addition, the computations are significantly complicated, since the coefficient α depends on the magnitude of capacitance C and on the range within whose limits the variations in Δe_ω occur.

This circumstance makes it mandatory, in the choice of parameters for the feedback loop in this case, to use expression (14), while keeping in mind the possibility of deviation of the actual duration from the computed one, these deviations being, as a rule, on the side of accelerating the transient response.

In particular, if $\alpha < 1$, $\beta = 0$, $\beta_a = 0$ and $R_a = 0$, then expression (15) takes the form:

$$\theta = \alpha\left(1 + \frac{CR_L}{\alpha\tau_v}\right).$$

When $CR_L \ll \alpha\tau_v$, $\theta \approx \alpha < 1$. This means that, by connecting a capacitance across the load, one can, in certain cases, obtain an acceleration of the transient response even when no flexible feedback is present.

As an illustration, we give the oscillograms of Fig. 7 (cf. Appendix 3). In taking off the oscillogram of Fig. 7b, the process occurred more rapidly, in spite of the fact that the load was shunted by a capacitance which was absent in the circuit corresponding to the oscillogram of Fig. 7a. Analogous conclusions, obtained by a different route, are contained in [8].

In the case of an inductive load (R_L, L_L), the analysis is complicated due to the raising of the orders of Eqs. (4), (5) and (6), and also due to the necessity of taking into account the switching in the diodes, the character of which, in its turn, depends on the capacitance. Since the expressions obtained are quite horrible, they are not included in the present work.

The use, with an inductive load, of the feedback with parameters chosen under the assumption that $L_L = 0$ gives satisfactory results. However, it is necessary to bear the following in mind with this. Even with no feedback present ($\beta = 0$ and $\beta_a = 0$) and with the proper choice of capacitance, autooscillations of the relaxation type may be observed. Figure 8 (cf. Appendix 3) shows the oscillogram of such autooscillations.

SUMMARY

1. By means of an integral quadratic estimate, one can choose the optimal parameters of the feedback path and avoid the errors which could arise if the parameters were chosen from the condition that a critical aperiodic transient response be obtained.

The increase of speed of the transient response due to the introduction of a flexible feedback will be greater when there is a rigid feedback present and for values of $\beta_a k_i = 2$ to 4.

2. When the load is shunted by a capacitance and there is no feedback present, there may be observed, in certain cases, an acceleration of the transient response and the generation of autooscillations.

APPENDIX 1.

Description of the Amplifier Tested

The amplifier cores were spun toroids of EN80Me (NII TVCh); $Q = 2 \times 0.325$ cm^2, $l = 0.13$ meter, $\mu_i = 62$ millihenries/meter, $\sigma = 0.98$, $\mu_c = 0.6$ millihenries/meter, $\nu = 1.38$. The amplifier windings have the following data: $W_\omega = 190$, $W_0 = 50$ and 200, $W_\beta = 185$, $\beta = 3.7$. The amplifier's internal impedance is $R_i = 3500$ ohm, its characteristic steepness was $S = 0.26$ and 1.05. The amplifier load was $R_L = 184$ ohm, $L_L = 25$ henries; the diodes were type DG-Ts25, $r_a = 3$ ohm.

APPENDIX 2.

The curves of Fig. 3 were obtained from the specimen described in Appendix 1. The voltages are expressed in average values and were measure by a rectifying instrument for $L_L = 0$.

Curves on Fig. 3	β	β_a	C, μf
1	—	—	—
2	3.7	10.3	4.7
3	—	—	1.25

APPENDIX 3.

The following parameters and modes of operation correspond to the oscillograms:

Figure number		β	β_a	C, μf	ΔI_c, ma	Initial magnetization	$\frac{\beta_a k_i}{m}$
4	a	—	—	—	80	20	1.07
	b	—	4.0	42.0	80	20	1.07
	a	3.7	—	—	10	Due to feedback	0
	b	3.7	3.7	285.0	10	Ditto	1.0
	c	3.7	4.3	60.3	10	" "	1.15
6	d	3.7	5.7	18.3	10	" "	1.5
	e	3.7	7.7	9.5	10	" "	2.05
	f	3.7	10.3	6.0	10	" "	2.75
7	a	—	—	—	20	10	0
	b	—	—	1.25	20	10	0
8		—	—	1.25	—	10	0

Note: On Fig. 6b, the magnitude of capacity was chosen arbitrarily; on Fig. 8, $L_L = 25$ henries.

APPENDIX 4.

The following parameters correspond to the curves of Fig. 5: $R_0 = 2.4$ ohm; $R_L = 185$ ohm, $R_L + 2r_a = 190$ ohm, $m = 0.95$, $n = 0.075$, $R_a = 4$ ohm, $1/(1 - \beta k_i) = 16$, $k_{i\beta} = 4.1$, $\tau_{\nu\beta} = 3.7$ millisecond, $\tau_{a\beta} = 6.4$ millisecond.

For curves 1 and 2, $r_a = 0$, $m = 1$, $n = 0$, $\tau_{a\beta} = 0$.

APPENDIX 5.

For $\alpha \neq 1$, Eqs. (4) and (5) take the form:

$$\Delta B_0 = \frac{\mu_c W_0}{l}(\Delta i_0 + \beta \Delta i_L + \beta_a \Delta i_a) + \\ + \frac{\alpha\nu}{\omega} \frac{R_L \Delta i_L + 2r_a(\Delta i_L + \Delta i_a) + W_\beta Q \frac{d(\Delta B_0)}{dt}}{W_\omega Q},$$

$$SR_i(\Delta i_0 + \beta \Delta i_L + \beta \Delta i_a) = \alpha R_L \Delta i_L + (R_i + 2\alpha r_a)(\Delta i_L + \Delta i_a) + \alpha W_\beta Q \frac{d(\Delta B_0)}{dt}.$$

LITERATURE CITED

[1] M. A. Rozenblat, Magnetic Amplifiers [In Russian], page 684 Sovetskoe Radio (1956).

[2] O. L. Aven, "Decreasing magnetic amplifier lags by the introduction of flexible feedback," [In Russian], Automation and Remote Control (USSR) 18, 2 (1957).*

[3] A. A. Krasovskii, "Integral estimates and the choice of parameters for automatic control systems," [In Russian]. In the book, Automatic Control Fundamentals [In Russian] (edited by V. V. Solodovnikov). Mashgiz (1954).

[4] L. V. Safris, "Magnetic amplifier design by linearized magnetization characteristics," [In Russian], Trudy Rostovskogo Instityta Inzhenerov Zh.-D. Transporta 20. Transzheldorizdat (1956).

[5] V. I. Molotkov, "The operation of a magnetic amplifier on a diode bridge with capacitive filters," In Russian]. Trudy Leningradskogo Politekhnichesogo Instuta Im. Kalinina 194 (1958).

[6] N. A. Kaluzhnikov, "On the design of a choke-controlled magnetic amplifier connected to a single-phase diode bridge," [In Russian]. Automation and Remote Control (USSR) 19, 3 (1958).*

[7] A. A. Krasovskii, "On the degree of stability of linear systems," [In Russian]. Trudy VVIA Im. Zhukovskogo (1948).

[8] V. I. Molotkov, "Investigation of the transient response in a magnetic amplifier with capacitive filters at the output of a rectifier bridge," [In Russian]. Nauchno-tekhnicheskii Informatsionnyi Byulleten' LPI Im. M. I. Kalinina 10. Radiofizika (1957).

Received November 3, 1959

* See English translation .

ON INCREASING TELEMETERING EFFECTIVENESS

N. V. Pozin

(Moscow)

Ways are considered for increasing the effectiveness of information transmission by telemetry on the basis of taking into account such spectral characteristics of the measured parameter as the dependency on frequency of oscillation of the amplitude of the oscillations and the measurement errors.

The paper presents a method of transmission with automatic changeover of the coding system. Estimates are made of the increase in speed of transmission and of the possible decrease in bandwidth.

Increasing the effectiveness of transmission allows one to decrease the bandwidth or to decrease the transmission time, and is attained, for example, on the basis of taking the properties of the communications into account in the design of the coding system. In information theory, Shannon has given a general definition which includes the statistical aspects of information transmission. The effectiveness of transmission is defined as the ratio of "the actual speed of transmission to the channel capacity" [1]. For judging practical transmission systems it is advantageous to use the concept of relative effectiveness (efficiency) η, where we understand by this the ratio of the maximum speed of information production by a given source, R_S, to the maximum speed of transmission of information analogous in its statistical and spectral properties by means of a coding system to be employed, R_C, $\eta = R_S/R_C$. Sometimes the increase in efficiency is conveniently estimated in the form of the gain (ϑ) in speed of transmission (R): $\vartheta_R\% = (1-(R_C/R_{C0}))\,100$, or $\vartheta_R = R_{C0}/R_C$, where R_{C0} and R_C are the original and the newly obtained values of speed of information transmission by the coding system. However, information-theoretical estimates (of transmission rate, for example) are ordinarily averages (over individual transmissions, over time) so that the use of them for comparing a number of transmission systems, telemetering systems in particular, is not always indicative. The specific feature of telemetry is that the lags connected with optimal coding are ordinarily inadmissible and, consequently, it is impossible to increase the average transmission rate by this method.* Therefore, the increase in efficiency is more frequently estimated directly in terms of the physical indicators of the transmission system, for example, as the gain in bandwidth when one switches from the initial bandwidth W_0 to a newly attained bandwidth W, i.e.,

$$\vartheta_W\% = \left(1-\frac{W}{W_0}\right)100\% \quad \text{or} \quad \vartheta_W = \frac{W_0}{W}.$$

1. Spectral Characteristics of the Parameter to be Telemetered

It is possible in many cases to increase the relative efficiency of information transformation by taking into account the spectral characteristics of this information in designing the coding system.** In the given case, we consider as spectral characteristics the dependence on frequency f of the amplitude of the oscillations A (f),

*To the author's knowledge, this property of telemetered messages was first formulated by V. A. Il'inoi.

**Here, the expression "coding system" is understood in the broad sense, i.e., as the transformer (cipherer, modulator) of one form of information-bearing signal to another form convenient for transmission along the communication channel.

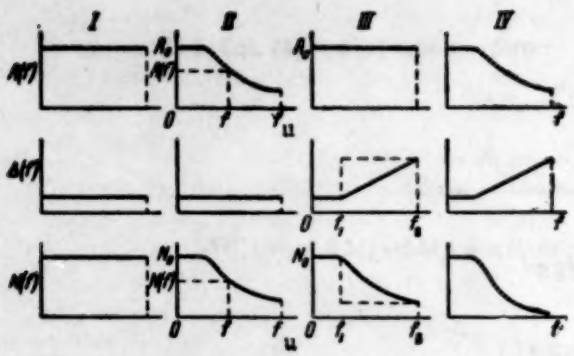

Fig. 1. Spectral characteristics of certain parameters subject to telemetering.

the admissible absolute error $\Delta(f)$ in the measurements and the number of gradations, $N(f) = A(f)/\Delta(f)$, necessary for transmission. These characteristics can be determined from an analysis, not only of the message source, but also of the receiver. Indeed, the properties of the receiver which render unnecessary the transmission of some elements or other of the messages must be taken into account in determining the characteristics of the messages to be transmitted. Thus, the spectral characteristics can be given to the input of the coding system.

Figure 1 shows the most typical spectral characteristics which are conventionally divided into four types. We note that the second type of characteristic corresponds to a decrease in the amplitude of the oscillations, and the third to a increase in admissible error, with increasing speed of variation (frequency of oscillation) of the telemetered parameter. The fourth case is a combination of the second and third cases.

Ordinarily, telemetry devices are designed in correspondence with characteristics of the first type. However, if the actual characteristics of the telemetered parameter are different from characteristics of the first type then, in many cases, it is disadvantageous to design a device for the maximum frequency of signal variation and the maximum number of gradations.

One method of matching channel (or coding system) and signal spectral characteristics consists of dividing the frequency band occupied by the signal into portions with characteristics close to uniform, and then designing several coding systems in correspondence with these portions [2, 3].

We propose another method of matching channel and signal characteristics for telemetry, this method consisting of the construction of an automatic switchover between coding systems, automatically varying the number of transmitted gradations, at the expense of varying the speed of operation, as a function of the speed of parameter variation.* Automatic switchover is essentially a change of a component of signal volume V** although, in many cases, for example in pulse-code modulation, the space itself, $V = 2fTN$, is changed.

2. Increasing Efficiency for Characteristics of the Second Type

We now consider the interval between pulse transmissions, corresponding to a parameter change with frequency f (Fig. 2). The parameter values x_i and x_{i+1}, separated in time by $1/2f$, can differ only within the limits of the zone $\pm A(f)$. The actual gain in information at each ordinate can reach the magnitude $\Delta I = \log [A(f)/\Delta_0] = \log N(f)$. Since the number of gradations which distinguish one transmission from another does not exceed $N(f) < N_0$ (Fig. 1,II), it is then advantageous to transmit, at each transmission, precisely that number $N(f)$ of gradations within whose limits lie the possible parameter variations at the given frequency. In telemetry, however, the transmission of just one change is, as a rule, inadmissible, since it is necessary to have correction indications to exclude the possible observational errors. This leads to a device of such a structure that the transmission of parameter variations is joined with the more infrequent transmission of absolute values.

The block schematic of a device in which matching of source and channel is implemented by means of several, two for example, coding systems can be the following. A slow-acting device with pulse-code modulation (PCM) transmits, with a long period, the accurate values of the parameter. By a separate channel are transmitted the rapid changes, for example, by the method of Δ-modulation or by one of its variants, distinguished by the fact that, in the absence of parameter variations, no signal is transmitted along this channel. Figure 3 shows the block schematic of a device which implements both these functions. Without going into details, we note that the reversible counter is very convenient in such code devices, and eases the isolation of deviation

*The character of such transmission has analogies in certain biological processes, for example, in certain human reactions.

**Such a possibility was stated by A. A. Kharkevich [4]. The definition of signal volume used here is somewhat different from that given by A. A. Kharkevich, which is closer to the established definition of speed of transmission.

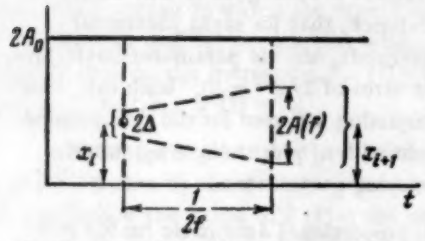

Fig. 2. Transition from parameter value x_i to x_{i+1} in the interval 1/2f.

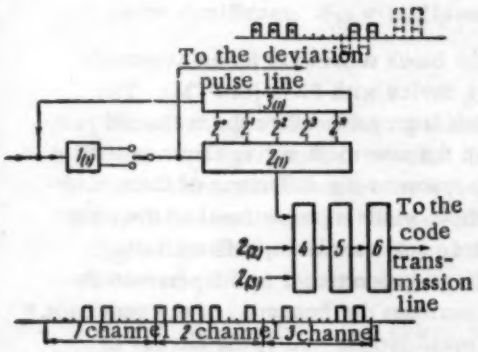

Fig. 3. Block schematic of a multichannel telemetry code device transmitting, as an additional feature, the rapid variations of one of the parameters by an individual channel. 1) the null-organ (the output is represented conventionally by contacts); 2) the binary reversible counter; 3) transforms code to voltage; 4) the channel switcher; 5) the repeater-index and 6 the inquiry switcher. The subscripts in parentheses are the channel numbers.

pulses whose values can differ by units. The necessity, in principle, of three channels for transmitting one telemetered quantity sharply decreases the desirability of such a method.

A more promising method of telemetering is the construction of automatically changed-over coding systems which transmit deviations and absolute parameter values over one individual channel. Of the possible variants, we mention the following code transmission: for the rapid variations of the parameter, there are two-pulse code transmissions of one of two combinations which assume the values + 1 or − 1 in the system of Δ-modulation.* For transmitting a constant component, when the parameter varies slowly, each such two-pulse transmission consists of one of two other combinations which correspond to 1 or 0 in the binary code.

3. Increasing Efficiency for Characteristics of the Third Type

Characteristics of the third type correspond to cases when the admissible telemetering error increases with increases in the frequency of oscillations (i.e., the speed of variation) of the parameter, and there is one and the same maximum value of these oscillations, i.e., $A(f)_{max} = A_0$. In telemetering devices which transmit such information it is also advantageous to use automatic changeover of coding systems.

Thus, for example, during relatively stable operation of a dispatcher-controlled energy system, one may record accurate data by telemetric means, process these data and then carry out the acts called for by these data. In the case of an abrupt variation in the controlled parameters, due, for example, to an accident, the dispatcher cannot judge the then occurring events with the same original accuracy, so that the transmission of telemetry data can be rougher in this case.

At the same time, it is important to the dispatcher to obtain these data with the least possible delay.

Automatic changeover telemetry devices can be implemented with any type of modulation. The most easily realized are telemetry devices with pulse-width, pulse-time and pulse-code modulation (PWM, PTM, PCM).

In the case of PCM, several methods of implementing discrete automatic tuning exist. One of these consists of the development of a coder which transmits, for slow parameter variations, code transmissions consisting of n_1 places of a binary code and, for rapid variations, shorter transmissions of $n_2 < n_1$ places of the code. A signal of the transition to the rapid, but rough, transmission might be, for example, the absence of a synchronizing pulse. The receiver must have a decoding device which discriminates the short and long transmissions and decodes

*The values of these Δ-transmissions can also differ from units, which corresponds to characteristics of the fourth type. Moreover, the elementary transmission may be, not two-, but multi-pulse (cf. section 3).

Multi-channel transmission with time separation is implemented by such a method. The specific feature here is the necessity of transmitting initially the first elements of the codes for all channels successively, then the second elements, etc.

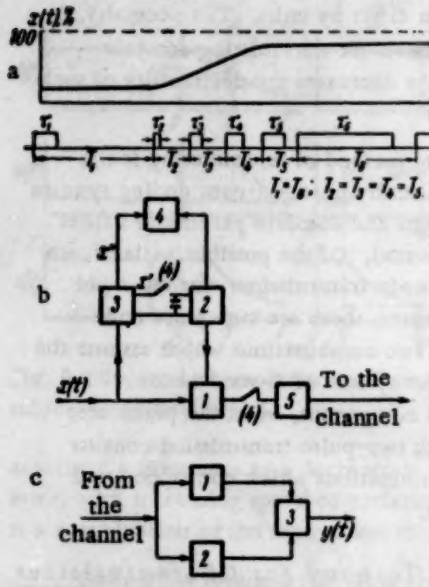

Fig. 4. Pulse-width telemetry device with automatic change-over (tuning). a) the cycle variation for abruptly changing parameters; b) the sending set: 1) the null organ; 2) the saw-tooth voltage generator; 3) the differentiating block; 4) the relay element; 5) the output pulse sharper; c) the receiver: 1) the demodulator of measurement time interval τ; 2) the demodulator of cycle time T; 3) the divider for τ/T.

them. Another possible variant is analogous to that described in section 2 with this difference, that for rapid parameter variations, the n_2 pulses transmit, not the parameter deviations, but the full scale, with an error of $1/2^{n2} - 2$). With this, two combinations in this transmission are used for the transmission of 1 or 0 in the binary code with n_1 places ($n_1 > n_2$) used in transmissions for slowly varying parameters.

With PWM or PTM, proportional automatic tuning is carried out. Figure 4a shows an example of the change in duration of the measurement cycle (transmission) T in the case of PWM for two velocity changes of the parameter (τ is the measurement time interval).

Figure 4b gives the block schematic for an automatic changeover transmitting device with PWM (or PTM). The transducer voltage, which is proportional to the measured parameter, is compared with the saw-tooth voltage from generator 2. Null-organ 1, which reacts to the difference of these voltages, develops pulses whose width is proportional to the transducer voltage. With parameter variations, differentiating block 3 generates a voltage (proportional to the parameter's first derivative) which increases the frequency of the generator.* So as not to distort the measurement cycle, the change in saw-tooth voltage generator frequency is carried out at the moment when this cycle terminates.

To exclude the delays in telemetry transmission which are possible with this, it can be provided that a sufficiently large second derivative of the parameter could interrupt a cycle already begun, after which a new cycle would commence with a new rate of increase of the saw-tooth voltage. An indication that an incomplete cycle had been interrupted can be transmitted, for example, by the simplest code transmission.

For the implementation of these functions there is a relay element, 4, in the scheme, whose executive organ is conventionally represented by contacts. The relay element is controlled by the differentiating block (the second derivative) and by the generator (at the moment of cycle termination). The invariance of the slope of the saw-tooth during the measuring cycles is provided by a memory device in the form of a condenser.

The receiving device (Fig. 4c) contains two demodulators, 1 and 2, for the separate demodulation of the time interval, proportional to parameter τ, and the current value of cycle time T. For obtaining indications, it is necessary to divide these quantities by means of special divider 3, for example, a ratiometer.

When the pulse transmission repetition rate is increased, there is a decrease in the telemetry accuracy of the code variant of the telemetry device due to a decrease in the number of transmitted places in the binary code, and a decrease in the pulse-width variant due to a decrease in the absolute value of the time deviations.

To estimate the gain in efficiency of such systems, we consider an example. Let the telemetry device with PWM transmit a parameter whose characteristics are given in Fig. 1, III.

As a preliminary, we note that a nonchangeover coding device, designed for upper frequency f_U and number of gradations N_0, must transmit, per second,

$$R_C = 2f_L \log_2 N_0 \frac{\text{bits}}{\text{sec}} \quad \text{or} \quad V = 2f_L N_0 \frac{\text{degrees}}{\text{sec}}$$

*The differentiating block operates on the modulus of the derivative, i.e., its output signal is independent of the sign of the parameter variation.

In our example, we take the case of a broken-line function $\Delta(f)$, corresponding to a PWM device with automatic tuning in which the repetition frequency of pulse transmissions is made proportional to the frequency of parameter oscillation, starting with some value, f_1, of this frequency. With this, the number of bits per second will depend on the frequency, $R_C(f) = 2f \log N(f)$, and the number of gradations transmitted per second does not change, i.e., $V(f) = V_0$.

In this somewhat idealized example, the rate at which the source generates information equals its speed of transmission by the coding system, $\eta = 1$. The gain in transmission rate is $\vartheta_R = R_{C0}/\max R_C = \log N_0/\log N(f_U)$, since the value of $R(f)$ is not constant in the range $0 < f < f_U$, and $\max R_C(f) = R_C(f_U) = 2f_U \log N(f_U)$.

We now determine the gain in bandwidth. If, with a nonchangeover device, the bandwidth in the channel is proportional to f_U then, in our case, this bandwidth is proportional to frequency f_1 with easily implemented conditions on the invariance on the edge steepness in the process of automatic changeover. The gain in bandwidth can be significant: $\vartheta_W = W_0/W = f_U/f_1$.

SUMMARY

The advantages of special measures directed towards increasing telemetry efficiency are defined, certainly, by a number of factors and, in the first instance, by the simplicity of their technological implementations. Together with such well-known methods for fuller use of communications channels as improving the quality of elements and components (stability of filters, generators, etc), increasing telemetering efficiency on the basis of taking the spectral characteristics of the measured parameter into account allows an additional manifold decrease in bandwidth or increase in tele-transmission speed.

To match the information source with the channel in telemetry, it is advantageous to use an automatically changed-over coding system which varies the number of transmitted gradations and the rate of telemetering as functions of the speed of parameter variation.

It is necessary to mention that although the basic types (second and third) of spectral characteristics considered contain decreasing functions of numbers of gradations on frequency, i.e., $N(f)$, nonetheless the difference in characteristic $A(f)$, and also $\Delta(f)$, determines the specific features of the transmissions systems for parameters of each type.

Whether some spectral characteristics of parameters or others are typical cannot be definitively decided today. However, it seems most probable that, in industrial telemetry, the most widespread will be the parameters with characteristics of the fourth type where, with increases in frequency, there are permitted, not only increases in the measurement errors, but also decreases in the amplitude of parameter oscillations.

LITERATURE CITED

[1] C. Shannon, "Statistical theory of electrical signal transmission," [Russian translation of English original], in collection of translations, Theory of Electric Signal Transmission in the Presence of Noise. [In Russian], (edited by N. A. Zheleznov). IL (1953).

[2] E. E. David and H. S. McDonald, A Bit-Squeezing Technique Applied to Speech-Signals. IRE, Conv. Rec. 4, No. 4 (1956).

[3] E. R. Kretzmer, Reduced-Alphabet Representation of Television Signal. IRE, Conv. Rec. 4, No. 4 (1956).

[4] A. A. Kharkovich, An Outline of the General Theory of Communicaton. [In Russian]. Cos. izdatelstvo techn.-teoret. lit-ry. (Moscow, 1955).

Received April 20, 1959

SWITCHING CIRCUIT OPERATION DURING TRANSITION PERIODS.

V. N. Roginskii

(Moscow)

A method is given for disclosing the possible disruptions in switching circuit operation during the periods of operation or release of the individual relays in the circuit, and recommendations are made for avoiding these disruptions.

In the course of switching circuit operation, the changes in individual relay states do not occur instantaneously and, during the "transition periods," it is possible that the contact circuits will be disturbed due to the fact that individual contacts of one and the same relay do not react (make or break) simultaneously. Thus, for example, if, at operation of a relay, a break contact is initially broken after which a make contact is closed, there will then exist a period during which both contacts will be open. If the contacts react in the reverse order, there will then be a period of time during which both contacts will be closed. Either one or the other condition might lead to a disruption of proper circuit operation since, in the first case, there might be a momentary open circuit and, in the second case, a false circuit might appear. This, in its turn, might lead to the incorrect operation or release of a relay or other device on which the given circuit operates.

The nonsimultaneous making and breaking of the individual contacts of one relay might be caused, on the one hand, by the design of the contact group, and can be systematic and, on the other hand, by the mechanical properties of the individual springs, and may be random. Analogous phenomena are also observed in cases when two or more relays in the circuit simultaneously change their state and, due to temporal separations of their operations or releasings, their contacts do not react simultaneously.

Switching theory [1, 2], at the beginning of its development, studied circuits only in their "static" states, and did not consider the transition periods. Recently, a number of works, [3, 4, 5] have appeared which consider individual aspects of switching circuit operation in the transition periods. In the present work, we present a general analytic method for accounting for switching circuit disruptions during the transition periods, and we consider methods for compensating these disruptions.

1. The Appearance of Switching Circuit Disruptions in the Transition Periods

The transition processes in switching circuits will be considered as the severing of mechanical connections between individual contacts of one relay, i.e., these contacts will be treated as contacts of different "ficticious" relays, analogously to what was done by Gr. Moisil and Gh. Ioanin [3].

First of all, we consider the case of disruptions which can appear in a switching circuit when the state of some relay X changes.

It can be shown that homologous contacts (all make, or all break) cannot disrupt circuit operation during the transition periods and, therefore, analysis is required only for circuits in which make and break contacts of relay X enter simultaneously.

For the analysis, we use the method of elimination of variables [6] by expressing the circuit $f(x, y, \ldots, z)$

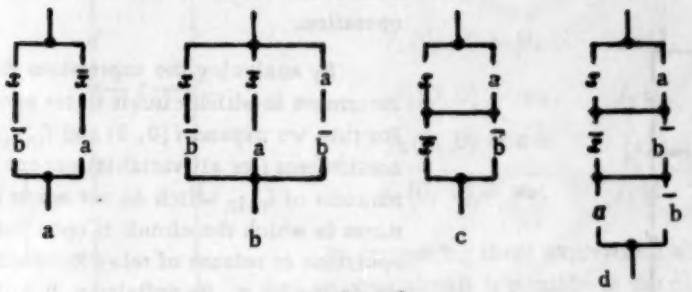

Fig. 1.

in terms of the contacts of X [1, 2]:

$$f(x, \bar{x}, y, \ldots, \bar{z}) = xf(1, 0, y, \ldots, \bar{z}) + \bar{x}f(0, 1, y, \ldots, \bar{z}). \quad (1)$$

Since we are only interested in contact $\underline{x}$, we shorten expression (1) to the following form:

$$f(x, \bar{x}) = xf(1.0) + \bar{x}f(0.1), \quad (2)$$

where f (1, 0) and f (0, 1) are functions which include all remaining variables but $\underline{x}$.

It can be shown that

$$f(1.0)f(0.1) = f_{\min} \quad (3)$$

is an exact lower bound of the function f (x, $\bar{x}$) since it contains only those subconstituents [2] (i.e., constituents of the circuit which consists of contacts of relays y,..., z) which enter into the circuit simultaneously in conjunction with $\underline{x}$ and with $\bar{x}$. On the other hand,

$$f(1.0) + f(0.1) = f_{\max} \quad (4)$$

is an exact upper bound, since it contains all the subconstituents entering into the circuit.

In the transition period, as was already stated, there is a severing of the mechanical connection between make and break contacts of one relay, X.

It is completely obvious that no disruptions in operation will be observed if, for all possible combinations of states of the individual contacts of relay X in the circuit, no new circuits appear and no existing circuits are disrupted, and this corresponds to the assertion that so new subconstituents must appear and no existing subconstituents must be expunged.

We now analyze the case when the circuit contains two nonhomologous contacts of relay X: make contact $\underline{x}$ and break contact $\bar{x}$, which we shall consider as contacts of the individual relays X_1 and X_2. Thus, we have implemented the transition from the circuit f (x, $\bar{x}$) to the circuit f (x_1, x_2). If the scheme is described by structural formulae or by the matrix of immediate conductions, then this replacement entails the transformation of these formulae or matrices.

As was already stated, circuit disruption is possible in the case when both contacts turn out to be simultaneously open or closed.

In the first case, for explaining the possibilities of momentary open circuits, we substitute the values $x_1 = x_2 = 0$ in the structural formula. If the expression thus obtained, f (0, 0), equals $f_{\min}$, this indicates that none of the circuits will be disrupted. If one obtains the inequality*

$$f(0.0) \leqslant f_{\min} = f(1.0)f(0.1), \quad (5)$$

*The symbol $f_1 \leq f_2$, or $f_2 \geq f_1$, (f_2 includes f_1) indicates that f_2 contains all the constituents which occur in f_1 and also some quantity of additional constituents [6, 7].

Fig. 2.

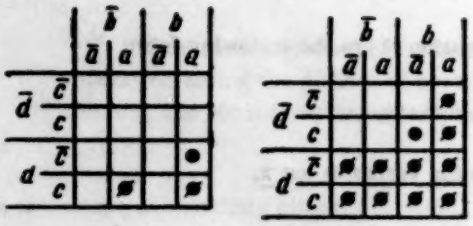

Fig. 3.

then this indicates that there can be disruptions in circuit operation.

By analyzing the expressions for f(0, 0) and $\underline{f}$, we can determine in which circuit states open circuit can occur. For this, we expand f(0, 0) and f_{min} in terms of the subconstituents (for all variables except $\underline{x}$), and those subconstituents of f_{min} which do not occur in f(0, 0) define the states in which the circuit is open during the periods of operation or release of relay X. Such "openable circuits" we denote by $\underline{g}$. By definition, it follows that

$$g = f_{min}\overline{f(0.0)}. \tag{6}$$

For finding the constituents of the "openable circuits" we can recommend the use of the coordinate method of writing Boolean functions [8], where one notes, by points for example, those quadrants which correspond to f_{min}, and then strikes out those quadrants which correspond to F (0, 0). The nonstricken quadrants with points correspond to those circuit states in which open circuits will appear.

In the second case, when both contacts are simultaneously closed, i.e., when $x_1 = x_2 = 1$, false circuits can appear if the following inequality is obtained:

$$f(1.1) \geqslant f_{max} = f(1.0) + f(0.1). \tag{7}$$

Those states in which false circuits can appear may be found by a method analogous to the one given previously if one notes by points those quadrants corresponding to f(1, 1), and then strikes out those of them which correspond to f_{max}.

The sum of the subconstituents of f(1, 1) which do not occur in f_{max} we call the "false circuits," and denote by φ, where

$$\varphi = \overline{f}_{max} f(1.1). \tag{8}$$

As an example, we consider the circuits shown in Fig. 1, for which we can write the structural formulae:

$$\begin{aligned} f_1(x, \bar{x}) &= x\bar{b} + \bar{x}a, & f_3(x, \bar{x}) &= (x + a)(\bar{x} + \bar{b}) \\ f_2(x, \bar{x}) &= x\bar{b} + \bar{x}a + a\bar{b}, & f_4(x, \bar{x}) &= (x + a)(\bar{x} + \bar{b})(a + \bar{b}). \end{aligned} \tag{9}$$

For all the circuits we get

$$f_{max} = ab. \tag{10}$$

$$f_{max} = a + \bar{b}. \tag{11}$$

By replacing $\underline{x}$ by x_1 and $\bar{x}$ by x_2, we get

$$\begin{aligned} f_1(x_1, x_2) &= x_1\bar{b} + x_2a, & f_3(x_1, x_2) &= (x_1 + a)(x_2 + \bar{b}) \\ f_2(x_1, x_2) &= x_1\bar{b} + x_2a + a\bar{b}, & f_4(x_1, x_2) &= (x_1 + a)(x_2 + \bar{b})(a + \bar{b}). \end{aligned} \tag{12}$$

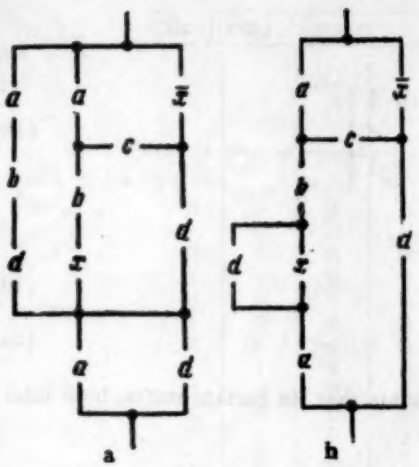

Fig. 4.

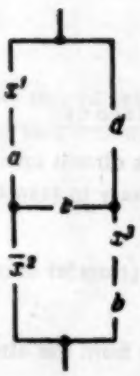

Fig. 5.

From whence,

$$\begin{aligned} f_1(0,\ 0) &= 0, & f_1(1,\ 1) &= a+\bar{b} \\ f_2(0,\ 0) &= a\bar{b}. & f_2(1,\ 1) &= a+\bar{b}, \\ f_3(0,\ 0) &= a\bar{b} & f_3(1,\ 1) &= 1, \\ f_4(0,\ 0) &= a\bar{b}; & f_4(1,\ 1) &= a+\bar{b}. \end{aligned} \tag{13}$$

By comparing these expressions with (10) and (11), we can see that an open circuit is possible in the circuit of Fig. 1a and a false circuit can appear in the circuit of Fig. 1c. There will be no disruption of operation during the transition periods of the circuits of Fig. 1b and d.

By comparing f_{min} and $f_1(0, 0)$ we obtain $g = a\bar{b}$ and, from a comparison of $f_2(1, 1)$ and f_{max}, we get $\varphi = a\bar{b}$, i.e., in the circuit of Fig. 1a an error is possible when relay A operates and relay B does not operate and, in the circuit of Fig. 1c, an error is possible when A does not operate and B does.

We note that, generally speaking, it is not possible in practice to replace $\underline{x}$ and $\bar{x}$ in the structural formulae first by x_1 and by x_2 and thereafter by zeroes or ones. It is possible to substitute ones or zeroes instead of $\underline{x}$ and $\bar{x}$ directly in the structural formulae.

For class N circuits such a substitution must be carried out in the structural matrix of immediate conductions (or in the corresponding determinant), after which one can expand the determinant.

We consider, for example, the circuit shown in Fig. 2, for which the structural matrix will be

$$M = \begin{pmatrix} 1 & 0 & a & \bar{x} \\ 0 & 1 & bx & d \\ a & bx & 1 & c \\ \bar{x} & d & c & 1 \end{pmatrix}. \tag{14}$$

The conductance between poles 1 and 2 is expressed by the determinant

$$f_{12} = \begin{vmatrix} 0 & a & \bar{x} \\ bx & 1 & c \\ d & c & 1 \end{vmatrix}. \tag{15}$$

From which we obtain

$$f(x,\ \bar{x}) = acd + abx + d\bar{x},$$

$$f(1,\ 0) = ab + acd, \quad f(0,\ 1) = d,$$

$$f_{min} = abd + acd, \tag{16}$$

$$f_{max} = d + ab. \tag{17}$$

For the transition period, we get from (15) that

$$f(0,\ 0) = \begin{vmatrix} 0 & a & 0 \\ 0 & 1 & c \\ d & c & 1 \end{vmatrix}, \quad f(1,\ 1) = \begin{vmatrix} 0 & a & 1 \\ b & 1 & c \\ d & c & 1 \end{vmatrix}. \tag{18}$$

From whence

$$f(0,\ 0) = acd \leqslant f_{\min} \tag{19}$$

$$f(1,\ 1) = d + ab + bc \geqslant f_{\max} \tag{20}$$

Consequently, in the transient periods in the given circuit it is possible that, in certain states, both false circuits and open circuits can occur.

Figure 3a shows the finding of the open circuit for the given case

$$g = ab\bar{c}d. \tag{21}$$

Figure 3b shows the finding of the false circuit

$$\varphi = \bar{a}bc\bar{d}. \tag{22}$$

2. Measures to Prevent Circuit Disruptions in the Transition Periods

If it is observed that the circuit disruptions in the transition periods can influence circuit operation as a whole, i.e., if they occur in states of the circuit which are used [2, 9], then it is necessary to take corresponding measures. Among these measures are:

a) the use of sequence fixing of the actions of the individual contacts of a relay (transfer contacts, or successively acting contacts);

b) changing the structure of the circuit so as to exclude nonhomologous contacts from the circuit;

c) applying lags, covering the duration of the circuit disruptions due to transition processes, to the relays on which the given circuit acts;

d) the introduction of covering circuits which would remain closed or open during operation or release time of the relay whose switched contacts are in the main circuit.

The first three measures require no special explanation, so we shall consider only the conditions which must be satisfied by a covering circuit.

For the case of momentary open circuits in the transition period, the covering circuit is connected in parallel to the circuit under investigation, and must include, as a minimum, the subconstituents of all states in which an open circuit is possible, and can also include all the remaining subconstituents entering into f_{min}, i.e., the shunting covering circuit $\mathcal{G}\delta$ must satisfy the inequality

$$g \leqslant \phi \leqslant f_{\min} \tag{23}$$

The solution of this inequality can be written in the form [6]

$$\phi = g + \omega f_{\min} = g + \frac{f_{\min}}{0} = f_{\min} \frac{\overline{f(0,\ 0)}}{1}, \tag{24}$$

where ω is an arbitrary circuit.

In the simplest cases, the covering circuit can be found from the well-known equivalence

$$xa + \bar{x}b = xa + \bar{x}b + ab, \tag{25}$$

to which corresponds, for example, the transition from the circuit of Fig. 1a to the circuit of Fig. 1b.

Circuit contact states				$x^1-x^2-x^3$				$x^4-x^5-x^6$			
a	b	c	d	$f(0,0,0)=dc$	$f(1,0,0)=a+bc$	$f(1,1,0)=0$	$f(1,1,1)=b(d+ac)$	$f(0,0,0)=dc$	$f(0,0,1)=d(b+c)$	$f(0,1,1)=bd$	$f(1,1,1)=b(d+ac)$
0	0	0	0								
1	0	0	0		+						
0	1	0	0								
1	1	0	0		+						
0	0	1	0								
1	0	1	0		+						
0	1	1	0								
1	1	1	0		+		+				+
0	0	0	1								
1	0	0	1		+						
0	1	0	1				+		+	+	+
1	1	0	1		+		+		+	+	+
0	0	1	1	+	+			+	+		
1	0	1	1	+	+			+	+		
0	1	1	1	+	+		+	+	+	+	+
1	1	1	1	+	+		+	+	+	+	+

Fig. 6.

In the second case, when false circuits appear, the covering circuit must annihilate them. For this, the covering circuit ψ must be connected in series with the circuit being investigated, and must satisfy the inequality

$$\bar{\varphi} \geqslant \psi \geqslant f_{\max}, \tag{26}$$

i.e.,

$$\psi = f_{\max} + \omega\bar{\varphi} = f_{\max} + \frac{\bar{\varphi}}{0} = f_{\max} + \frac{\overline{f(1.1)}}{0}. \tag{27}$$

In the simplest case, one can use the equivalence

$$(x+a)(\bar{x}+b) = (x+a)(\bar{x}+b)(a+b), \tag{28}$$

to which corresponds the transition for the circuit of Fig. 1c to the circuit of Fig. 1d.

For the circuit of Fig. 2, we obtain from (16), (17), (21) and (22)

$$\varphi = abc\bar{d} + \frac{abd}{0} + \frac{acd}{0}, \quad \psi = d + ab + \\ + \frac{a+\bar{b}+\bar{c}+d}{0}.$$

In this case, the simplest covering circuits are

$$\varphi = abd, \quad \psi = d + a.$$

The corresponding circuit is shown in Fig. 4a and, in Fig. 4b, the same circuit, transformed by the general methods of contact circuits.

3. Dynamic Disruption

If there are make and break contacts of some relay X in a circuit, where the total number of contacts of this relay in the circuit is greater than two (i.e., there are several homologous contacts) then, together with the disruptions of which we spoke above, additional pulses can appear in the circuit during the transition periods [4]. Thus, for example, we consider the circuit of Fig. 5, where x^1, x^2 and x^3 are the numbered contacts of relay X. In the case when relays A, B and C operate (D does not operate), the circuit is closed when relay X operates.

However, if the contacts so react that initially contact x^1 closes, then contact x^2 breaks and only then does contact x^3 make, the circuit will initially close, then open and then close again, i.e., an additional pulse appears. One may easily convince oneself that such a pulse does not appear for any other sequence of contact reactions when relay X operates.

The disclosure of such disruptions can be done by the method described in this paper by considering each of the contacts of relay X as the contact of an independent ficticious relay and by investigating the possible sequences in the action of these relays. It can be shown that there can be disruptions only in the cases when between the reactions of two homologous contacts (make contacts, for example) there react one or several contacts of the other type (break contacts). Thus, for the circuit of Fig. 5, "dynamic" disruptions are possible only for the sequences $x^1 - x^2 - x^3$ or $x^3 - x^2 - x^1$. In other words, if the circuit is written in the form $f(x^1, x^2, x^3)$ then, in the first case, the sequence of states $f(0, 0, 0) - f(1, 0, 0) - f(1, 1, 0) - f(1, 1, 1)$ is chosen and, in the second case, the sequence $f(0, 0, 0) - f(0, 0, 1) - f(0, 1, 1) - f(1, 1, 1)$.

To disclose those states in which, for a given sequence of contact reactions, the appearance of "dynamic disruptions" is possible, one can use, for example, a table, constructed for each sequence, of circuit states corresponding to the individual subconstituents. Thus, Fig. 6 shows two such tables for the sequences of reactions $x^1 - x^2 - x^3$ and $x^3 - x^2 - x^1$. Here, the blank spaces show that the circuit corresponding to the given constituents will be open, and for the plusses will be closed, at each cycle engendered by the making or breaking of contacts of relay X. The sequence $- + - + (\text{or } + - + -)$ shows that, for the states of the circuit elements corresponding to the given constituent, an additional pulse is generated upon the closing (opening) of the circuit in the process of operating or releasing relay X. For the circuit of Fig. 5, $abc\bar{d}$ and $ab\bar{c}d$ are such states. In a number of other cases momentary closures and breaks are possible, but these are brought to light by an investigation carried out in accordance with section 1 of this paper.

If, for some reason, it is impossible to exclude some of the contacts of relay X from the circuit when there are "dynamic" disruptions, one can use covering circuits, analogous to those previously discussed.

4. Simultaneous Change of State of Several Relays

If there are cycles in the circuit during which several relays are operated or released simultaneously then the contacts of these relays may react, not simultaneously, but with some dispersion in time, determined not only by the properties of the contacts themselves, but also by the times of operation or release of the individual relays.

Analysis of such circuits and the disclosure of the possible appearances of disruptions can be carried out by the same methods as for contacts of one relay.

In the cases where one observes the possibility of disruptions of operation due to the nonsimultaneous change of state of different relays, one may, in addition to the methods mentioned above, introduce a delay in the relays so that the sequence of their operations is fixed, or one may duplicate the contacts of one relay by the contacts of another relay.

LITERATURE CITED

[1] M. A. Gavrilov, Switching Circuit Theory [In Russian], Izdatelstvo AN SSSR (1950).

[2] V. N. Roginskii and A. D. Kharkevich, Switching Circuits in Telephony [In Russian], Svyaz'izdat (1955).

[3] Gr. Moisil and Gh. Ioanin, Asupra functionaric schemelor cu butoni reali, Bul. stiitific sectiunea de stiinte matematice si fizice V. 7, No. 1 (1955).

[4] D. A. Huffman, The design and use of hazard-free switching networks, J. Assoc. Comput. Machinery vol. 4, No. 1 (1957).

[5] Ya. L. Mekler, "Transition periods in switching circuits," [In Russian], Automation and Remote Control (USSR) 18, 1 (1957).*

[6] L. Kutyura, The Algebra of Logic [In Russian], Izdatelstvo Matezis (Odessa, 1909).

*See English translation.

[7] V. N. Roginskii," Equivalent transformations of switching circuits," [In Russian]. Sb. Nauchnykh Rabot Po Provodnoi Svyazi, Vyp. 6. Izdatelstvo AN SSSR (1957). Collected Scientific Papers on Conductor Communications, vol. 6 [in Russian].

[8] O. Plechl, Elektromechanische Schaltungen und Schultgerute. Vienna (1956).

[9] V. N. Roginskii, "Taking unused states into account in the synthesis of switching circuits," [In Russian]. Automation and Remote Control (USSR) 15, 3 (1954).

Received January 6, 1959

ON ONE METHOD OF ANALYZING COMPLEX SWITCHING CIRCUITS*

V. F. D'yachenko

(Moscow)

A method is presented for the analysis of switching circuits which is based on the finding of the intermediate relay windings and executive elements carrying current, and their interconnections, for each state of the circuit's contacts.

Two-terminal, class P and N switching circuits with multiwinding relays and parametric dependencies are considered.

1. Introduction

The problem of analyzing switching circuits amounts to the determination, for nonsequential circuits, of the state of the executive circuit as a function of the states of the receiving elements and, for sequential circuits, the determination as well of the sequences of action of the intermediate and executive elements. For this analysis one must have, in addition to the structural scheme, data on the relays and other elements (resistances, number of turns, ampere-turns for operation and release, etc) which enter into the circuit. For multi-winding relays it is necessary to have data on the interactions of the individual windings.

The existing analytical methods [1-5] are based on the finding of the contact circuits which act on each intermediate and executive element of the switching circuit. Thus, for example, for a class P switching circuit [1], the structural formula

$$F = f(a_1, a_2, \ldots, a_k, x_1, x_2, \ldots, x_n, X_1, X_2, \ldots, X_n). \tag{1}$$

where X_j is a symbol for the reacting organ of an intermediate or executive element ($j = 1, 2, \ldots, n$), $a_1, a_2, \ldots, a_k$ are contacts of the receiving relays, $x_1, x_2, \ldots, x_n$ are contacts of the intermediate relays of $f_j = f_j(a_1, a_2, \ldots a_k, x_1, x_2, \ldots, x_n)$ is the contact circuit acting on element X_j, must be transformed to the form

$$E = f_1 X_1 + f_2 X_2 + \cdots + f_n X_n, \tag{2}$$

i.e., must go over to the group of circuits of the form $f_j X_j$.

For complex circuits, however, it is necessary in many cases to have recourse to additional graphical transformation of the circuits prior to separating out, by the existing methods, the contact circuits acting on each intermediate or executive element. For nonplanar circuits, and sometimes for circuits with parametric dependencies, it turns out to be necessary to use the existing methods to isolate the circuits acting on the individual elements of the circuit. The analysis is still further complicated for circuits with multi-winding relays. For example, the general form of the structural formula for a class P circuit with multi-winding relays is written as:

$$F = f(a_1, a_2, \ldots, a_k, x_1, x_2, \ldots, x_n, X_1^1, X_1^2, \ldots, X_2^1, X_2^2, \ldots, X_n^1, X_n^2, \ldots). \tag{1a}$$

*Presented at a joint seminar of the Laboratory for Wire Communications of the AN SSSR and the Telephony Dept. of the MEIC on January 15, 1958.

In individual cases one can go from this formula to the formula

$$F = f_{11} X_1^1 + f_{12} X_1^2 + \cdots + f_{21} X_2^1 + f_{22} X_2^2 + \cdots + f_{n1} X_n^1 + f_{n2} X_n^2 + \ldots, \qquad (2a)$$

where X_j^i (j = 1, 2,..., n; i = 1, 2, 3,...) is the symbol for the i'th winding of element $\underline{j}$ and $f_{ji} = f_{ji}(a_1, a_2, \ldots, a_k, x_1, x_2, \ldots, x_n)$ is the contact circuit acting on element X_j^i.

In this case, the action of multi-winding relay X_j depends not only on whether or not the circuit (f_{ji}) for some winding (X_j^i) is open or closed, but also on the interaction of the magnetic fluxes induced by the individual windings. In such cases, one should have recourse to the electrical computations of the currents flowing through the windings and of the magnetic fluxes induced by them.

2. Analysis of Nonsequential Class P Circuits

In the present work we consider two-terminal switching circuits. In order to determine whether or not an intermediate relay or executive element of the circuit operates, we must find, for each state of the circuit's contacts, the windings of the intermediate relays and executive elements through which current flows, and then determine the interconnection of the windings in the circuit.

For the analysis of class P switching circuits, we write the circuit in the form of structural formulae (1), (1a). With this, we adopt the convention that the symbol for a winding in which the current flows from end to beginning is written with a tilde (for example, $\tilde{X}_5^2$) while the symbol for a winding in which the current flows from beginning to end has no tilde (for example, X_5^1). Thereafter, instead of the contact symbols, we substitute the values of their structural conductances (1 for a make contact and 0 for a break contact) for a given state of the elements, and we transform the expression obtained, bearing in mind that

$$\begin{aligned} 0 \cdot W &= 0, \\ 0 + W &= W, \\ 1 \cdot W &= W, \\ 1 + W &= 1, \end{aligned} \qquad (3)$$

where W is the symbol for a relay winding or an impedance. After this, the structural formulae can assume one of the following forms.

1. F = 0. This says that the entire circuit between terminals of the battery is open, so that no current flows through any circuit element.

2. F = 1. This means that terminals are short-circuited, and that no current flows through the relay windings.

3. $F = f(X_1^1, X_1^2, \ldots, X_2^1, X_2^2, \ldots, X_n^1, X_n^2, \ldots, R)$. This asserts that the relay windings and impedances which correspond to the symbols actually in the formula carry current, where the sumbols +, and brackets (or parentheses) joining the symbols of these elements indicate that these elements are interconnected. By knowing the relationships of the parameters of these elements, we can determine which relays will operate for a given state of the circuit's contacts. For this, it is necessary to take into account, for multi-winding relays, the total action of all the current-bearing windings. A multi-winding relay will operate (given the condition that each winding induces the same number of ampere turns) if unequal numbers of fundamental windings (where the current flows from beginning to end) and counteracting windings (where the current flows from end to beginning) are carrying current, and will not operate if equal numbers of basic and counteracting windings are carrying current, even though the relay as a whole is carrying current.

If the interconnections between the circuit elements are such that the switching in of one element completely upsets the operation of another, i.e., the conditions of order of conductance are satisfied,* we may then

*The concept of order of conductance characterizes the action of some circuit element (an impedance, a relay winding), with conductance G, on the operation of a relay A. It is considered that, if conductance G is connected in series with relay A and so decreases the winding current that the relay neither operates nor holds but, with the two in parallel, relay operation is not disturbed, then the order of conductance of G is less than the order of conductance of A ($G_{<A}$). The concept of a higher order of conductance ($G_{>A}$) is introduced in an analogous fashion.

[Footnote continued on following page]

use the additional transformations for analysis of the switching circuit:

$$A_{<b} + A_{>A} = {}^{B}{}_{>A}, \tag{4}$$

$$A_{<b} \cdot {}^{B}{}_{>A} = A_{<b}.$$

If the interconnections between circuit elements are otherwise, and do not satisfy the conditions of order of conductance, we can then introduce other rules. Thus, for example, if impedance R, when connected in series with relay A, results in the latter not operating, but holding, then impedance R can be replaced by a contact, a, of relay A. If a parallel connection of relay A and impedance R causes A, not to operate, but to hold, then impedance R can be replaced by a contact a.

3. Analysis of Class N Circuits

In the analysis of a class N switching circuit [1], its structure is written in the form of a structural matrix [8, 9]

$$M = \| \alpha_{i,\,j} \| \quad (i,\ j = 1,\ 2,\ 3,\ \ldots,\ q) \tag{5}$$

where $\alpha_{i,j}$ is the directed path of current flow from node i to node j, with the convention that $\alpha_{i,\,i} = 1$. The relay winding symbols which enter into the expressions $\alpha_{i,\,j}$ must include a tilde if the beginning of the winding is connected to node j, and must be without a tilde if the connection is to node i. For multi-winding relays, each winding must necessarily bear its ordinal number, which is written in the form of a superscript.

If there are no rectifier elements in the circuit, the matrix will be symmetric with respect to the principal diagonal, except that the winding symbols on opposite sides of the diagonal will be differentiated by their bearing a tilde or not, i.e., if, for example, $\alpha_{i,\,j} = aX^1$, then $\alpha_{j,\,i} = a\widetilde{X}^1$, and if $\alpha_{i,\,j} = aY + X^2$, then $\alpha_{j,\,i} = a\widetilde{Y} + X^2$.

For the analysis of switching circuits with diode elements, we can consider that the impedance of a diode element in the conducting direction equals 0, and in the cut-off direction, equals ∞. Therefore, in setting up the structural matrix, a diode element is replaced by 1 if the direction of the current flow path coincides with the diode's conducting direction, and by a 0 for the reverse direction. In this case, the matrix will be non-symmetrical.

For example, for the circuit of Fig. 1, the matrix is written in the form:

$$M = \begin{Vmatrix} 1 & 0 & aX^1 & b + \bar{d}R_2 \\ 0 & 1 & \widetilde{Z} & R_1 + d\widetilde{Y} \\ aX^1 & Z & 1 & X^2 \\ b + \bar{d}R_2 & R_1 + dY & \widetilde{X}^2 & 1 \end{Vmatrix}. \tag{6}$$

For the circuit of Fig. 2, the matrix assumes the form:

$$M = \begin{Vmatrix} 1 & abZ & aB^1 & i \\ ab\widetilde{Z} & 1 & bR & \widetilde{A}^1 \\ a\widetilde{B}^1 & bR & 1 & (1 + \widetilde{B}^2)\,A^2 \\ i & A^1 & \widetilde{A}^2(B^2 + 0) & 1 \end{Vmatrix} \tag{7}$$

To find, from the matrix, all possible current-conduction paths in a circuit, it is necessary to strike out

*[Footnote continued]. by being applied to the case when a parallel connection of G and A causes relay A to be shunted and its operation disturbed. In the case when neither a parallel nor a series connection of conductance G and relay A affect the latter's operation, one considers that they have identical orders of conductance ($G = A$). For more details on this, cf., [6, 7].

In theory, the determination of the orders of conductance of circuit elements (impedances, relay windings) can be carried out prior to circuit structural analysis, and independently of it.

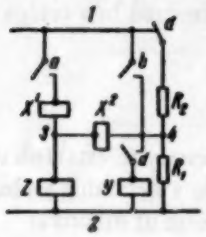

Fig. 1.

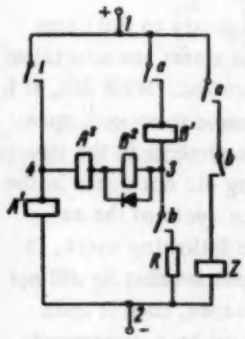

Fig. 2.

the column whose ordinal number coincides with the number of the node connected to the battery's plus terminal, and the row whose ordinal number coincides with the number of the node joined to the battery's minus terminal. (Such nodes are sometimes called the circuit's terminals). We expand the determinant obtained by the rules for expanding Boolean determinants [8, 9], i.e., we take all terms with a + sign, and take into account that, for contacts, $a \cdot a = a$ and $a \cdot \bar{a} = 0$. The expansion of Boolean determinants gives, not only the symbolic expressions for the nonself-intersecting current-conducting paths, but also the different redundant terms, which correspond to sequences of nodes which do not enter into the current-conducting paths. These terms are easily removable. In addition, in expanding a determinant which contains relay windings and elements of finite conductance, one must take the following transformations into account:

$$X_j^i \cdot \tilde{X}_j^i = 0, \tag{8}$$

$$F_1 + F_1 \cdot F_2 = F_1, \tag{9}$$

where F_1 and F_2 are arbitrary two-terminal circuits (in which, in addition to elements with finite conductances, relay contacts may also enter).

By thus expanding the determinant, we obtain an expression whose terms represent the symbolic expressions of the elementary nonself-intersecting current-conducting paths in the circuit between the + and − terminals of the battery. Further transformations are carried out in the same way as for the structural formulae of class P circuits. After substitution of the values of the structural conductances of the contacts for a given state of the circuit in the formula, one must transform the resulting expression so that, for each winding and each resistance, there would remain just one symbol in the formula, after which one can analyze relay operation as was done above. If it is impossible to lead the formula to a form where there is just one symbol for each winding, this indicates that the elements are connected in a bridge circuit. In this case, additional quantitative investigation is necessary.

We consider the switching circuit shown in Fig. 1, where it is known of the elements that all relay windings have the same order of conductance and, moreover $R_1 < YX^1X^2Z$. The matrix for the immediate conductances of this circuit was given above, (6). By striking from this matrix the first column, corresponding to the battery's plus terminal, and the second row, corresponding to the battery's minus terminal, we obtain a determinant whose expansion gives the elementary current-conducting paths from battery + to battery − :

$$F(a, b, d) = \begin{vmatrix} 0 & aX^1 & b + \bar{d}R_2 \\ Z & 1 & X^2 \\ R_1 + dY & \tilde{X}^2 & 1 \end{vmatrix} =$$

$$= aX^1X^2R_1 + adX^1X^2Y + bZ\tilde{X}^2 + \bar{d}R_2Z\tilde{X}^2 +$$

$$+ bR_1 + bdY + \bar{d}R_1R_2 + aX^1Z. \tag{10}$$

We now determine the state of the circuit's relays if, for example, receiving relays A and D operate, and relay B does not operate:

$$F(1, 0, 1) = X^1X^2R_1 + X^1X^2Y + X^1Z =$$
$$= X^1[X^2(R_{1<Y} + Y) + Z] = X^1[X^2Y + Z].$$

From this expression we can draw the conclusion that relays X, Y and Z will operate. If only receiving relay A operates, we get

$$F(1, 0, 0) = X^1X^2R_2 + R_2Z\tilde{X}^2 + R_1R_2 + X^1Z =$$
$$= X^1(X^2R_2 + Z) + R(Z\tilde{X}^2 + R_1).$$

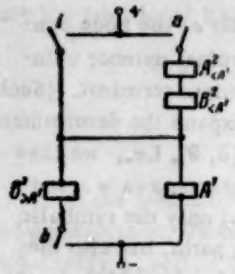

Fig. 3.

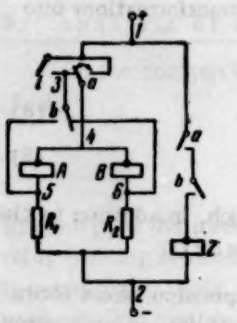

Fig. 4.

It is impossible to so transform this expression that there would be only one symbol for winding X^2 and, consequently, this winding is connected in a bridge circuit.

4. Analysis of Sequential Circuits

For the analysis of sequential circuits, it becomes necessary to establish the temporal sequence of circuit element actions, i.e., to obtain a switching table. For this, one first of all writes the circuit's structure in the form of structural formulae (1), (1a) (for class P circuits) or in the form of a structural matrix as in (5) (for class N circuits). It is also necessary to know which circuit elements are receiving, and which intermediate and executive.

The analysis of sequential circuits is carried out analogously to the cases already considered, with the sole difference that the element states are now taken in the same order as they occur in the process of circuit operation. With this, if it be assumed that intermediate relay X_j operated ($x_j = 1$) in some state and, upon substitution of the contacts' structural conductances for their symbols in the structural formula, it turned out that the symbol for this relay's winding did not occur in the right member, this would indicate that the given state (given cycle of the sequence) was unstable, and that it was necessary to go to the following cycle, in which relay X_j must not operate. Analogously, if it were assumed that X_j did not operate ($x_j = 0$), but the conditions for its operation were created, then it must operate at the following cycle. In the stable cycles there must be a correspondence between the states of the relays and the current flow through their windings. In this case, a transition to the following cycle must occur as the result of a change in an external stimulus. On the basis of such an investigation one constructs a switching table. If the sequence of external stimuli is unknown (the operation of the receiving elements), it is recommended that analysis of the sequential circuit operation begin with the state when all the receiving elements are nonoperated. This will be the zero cycle, which may be either stable or unstable.

In conclusion, we consider three examples of the analysis of the double pulse divider circuits shown on Figs. 2, 3 and 4, where I is the receiving (pulsed) relay, A and B are intermediate relays, and Z is the executive relay.

In the circuit of Fig. 2, the relay parameters and the resistance are so chosen that parallel or series connection of the windings of relays A, B, Z and resistance R does not disturb the operation of these relays. Relay A has two windings and, when current flows through both windings in different directions, relay A releases.

By striking out the first column and second row in the matrix, (7), for this circuit, we get

$$F = \begin{vmatrix} abZ & aB^1 & i \\ bR & 1 & A^2 \\ A^1 & \tilde{A}^2 B^2 & 1 \end{vmatrix} \tag{11}$$

By expanding this determinant, we obtain

$$F(i,\ a,\ b) = abZ + iA^1 + abB^1R + aA^1A^2B^1 + ib\tilde{A}^2B^2R =$$
$$= abZ + i(A^1 + b\tilde{A}^2B^2R) + a(bB^1R + A^1A^2B^1). \tag{12}$$

We assume that, in the initial state, all relays were nonoperated, i.e.,

$$F_0(0,\ 0,\ 0) = 0.$$

This will be a stable state of the circuit. To set the circuit in action, it is necessary that there be an external stimulus, i.e., it is necessary that receiving relay I operate. After this occurs, we shall have, in the first cycle

$$F_1(1,\ 0,\ 0) = A^1.$$

TABLE 1

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
I		—	—			—	—			—	—			—	—	
A		→	+	+	+	⊣				→	+	+	+	⊣		
B				→	+	+	+	⊣				→	+	+	+	⊣
A^1		·	⊙	⊙	⊙	·	·			·	⊙	⊙	⊙	·	·	
A^2				·	⊙	·	·					·	⊙	·	·	
B^1				·	⊙	⊙						·	⊙	⊙		
B^2					·	⊙	⊙						·	⊙	⊙	
Z					→	+	⊣						→	+	⊣	

This state is unstable, since there is no correspondence between the relay state (a = 0) and the flow of current through its winding. Relay A operates, and the circuit transfers to the following stable state:

$$F_2(1,\ 1,\ 0) = A^1 + A^1 A^2 B^1 = A^1.$$

To transfer to the following state, receiving relay I must release. Then we get

$$F_3(0,\ 1,\ 0) = A^1 A^2 B^1.$$

In this unstable state, relays A and B operate, with which

$$F_4(0,\ 1,\ 1) = Z + B^1 R + A^1 A^2 B^1 = Z + B^1 (R + A^1 A^2).$$

At this cycle we have a source for executive relay Z but the circuit state will be unstable. With the application of an external stimuli, receiving relay I again operates, with which

$$F_5(1,\ 1,\ 1) = Z + A^1 + \widetilde{A}^2 B^2 R + B^1 R + A^1 A^2 B^1 = Z + A^1 + (\widetilde{A}^2 B^2 + B^1) R.$$

It is clear from this expression that the currents in the windings of relay A will flow in opposite direction, so relay A must release, i.e., the circuit transfers to the following stable state:

$$F_6(1,\ 0,\ 1) = A^1 + \widetilde{A}^2 B^2 R.$$

After release of relay I we get

$$F_7(0,\ 0,\ 1) = 0.$$

Relay B releases, and the circuit returns to its initial state:

$$F_8(0,\ 0,\ 0) = 0.$$

If an external stimulus is again applied to the receiving relay, the sequence of circuit operation is repeated. On the basis of the analysis just given, we can construct the switching table (Table 1).

In the circuit of Fig. 3, the elements' parameters are so chosen that the order of conductance of windings A^2 and B^2 is less, and the order of conductance of winding B^1 is greater, than the order of conductance of winding A^1, i.e., $(A^2B^2)_{<A^1}$ and $B^1 > A^1$. Since this is a class P circuit, we immediately write its structural formula:

$$F(i,\ a,\ b) = [i + a\,(A^2 B^2)_{<A^1}]\,(bB^1_{>A^1} + A^1).$$

TABLE 2

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
I		—	—			—	—			—	—			—	—	
A		→	+	+	+	⊣				→	+	+	+	⊣		
B				→	+	+	+	⊣				→	+	+	+	⊣
A^1		·	⊙	·	·	·	·			·	⊙	·	·	·	·	
A^2				⊙	⊙							⊙	⊙			
B^1				·	·	⊙	⊙					·	·	⊙	⊙	
B^2				·	⊙							·	⊙			

TABLE 3

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
I		—	—			—	—			—	—			—	—	
A		→	+	+	+	⊣				→	+	+	+	⊣		
B				→	+	+	+	⊣				→	+	+	+	⊣
Z					→	+	⊣						→	+	⊣	

By reasoning as before, we obtain

$$F_0(0, 0, 0) = 0,$$
$$F_1(1, 0, 0) = A^1,$$
$$F_2(1, 1, 0) = A^1,$$
$$F_3(0, 1, 0) = (A^2B^2)_{<A^1} \cdot A^1 = A^2B^2,$$

$$F_4(0, 1, 1) = (A^2B^2)_{<A^1}(B^1_{<A^1} + A^1) = (A^2B^2)_{<A^1} \cdot A^1 = A^2B^2,$$

$$F_5(1, 1, 1) = B^1_{>A^1} + A^1 = B^1,$$
$$F_6(1, 0, 1) = B^1_{>A^1} + A^1 = B^1,$$
$$F_7(0, 0, 1) = 0,$$
$$F_8(0, 0, 0) = 0.$$

The circuit has returned to the initial state. On the basis of these results, we construct the switching table (Table 2).*

In the circuit of Fig. 4, the relay parameters and the impedances are so chosen that parallel or series connections of windings A, B, Z and resistors R_1, R_2 and R_3 does not disturb the operation of these relays.

The structural matrix for this circuit has the form:

$$M = \left\| \begin{array}{cccccc} 1 & abZ & i & a & 0 & 0 \\ ab\widetilde{Z} & 1 & 0 & 0 & R_1 & R_2 \\ i & 0 & 1 & \overline{a} & b & \overline{b} \\ a & 0 & \overline{a} & 1 & A & B \\ 0 & R_1 & b & \widetilde{A} & 1 & 0 \\ 0 & R_2 & \overline{b} & \widetilde{B} & 0 & 1 \end{array} \right\|. \qquad (13)$$

*Note. For clarity, we have written, in the lower portions of Tables 1 and 2, the relay windings, on the left, and, in the columns corresponding to the cycles of circuit operation, we have noted by dots those windings carrying current at the given cycle and, by circled dots, the windings through which current sufficient for relay operation flows.

We strike out the first column and second row of matrix (13), to obtain the determinant:

$$F = \begin{vmatrix} abZ & i & a & 0 & 0 \\ 0 & 1 & \bar{a} & b & \bar{b} \\ 0 & \bar{a} & 1 & A & B \\ R_1 & b & \tilde{A} & 1 & 0 \\ R_2 & \bar{b} & \tilde{B} & 0 & 1 \end{vmatrix} \tag{14}$$

By expanding this determinant, we get

$$F(i, a, b) = abZ + aAR_1 + aBR_2 + iaAR_1 + \bar{i}aBR_2 + \\ + ibR_1 + ibR_2\tilde{A}B + \bar{i}\bar{b}R_1\tilde{A}\tilde{B} + \bar{i}\bar{b}R_2 \tag{15}$$

Reasoning as before, we obtain

$$F_0(0, 0, 0) = 0$$
$$F_1(1, 0, 0) = AR_1 + R_2$$
$$F_2(1, 1, 0) = AR_1 + R_2,$$
$$F_3(0, 1, 0) = AR_1 + BR_2,$$
$$F_4(0, 1, 1) = Z + AR_1 + BR_2,$$
$$F_5(1, 1, 1) = Z + R_1 + BR_2,$$
$$F_6(1, 0, 1) = R_1 + BR_2,$$
$$F_7(0, 0, 1) = 0,$$
$$F_8(0, 0, 0) = 0.$$

The circuit has returned to the initial state. Table 3 was constructed on the basis of the analysis given.

The method considered for analyzing mixed switching circuits broadens our capabilities of analyzing the operation of a more general class of circuits which are widely used in contemporary devices of automation, remote control and communications.

LITERATURE CITED

[1] M. A. Gavrilov, Switching Circuit Theory. [In Russian]. Izdatelstvo AN SSSR (1950).

[2] V. N. Roginskii and A. D. Kharkevich, Switching Circuits in Telephony. [In Russian]. Xvyas'izdat (1955).

[3] M. A. Gavrilov, "Isolation, in switching circuits, of circuits acting on a given element," [In Russian]. Dokl. AN SSSR 87, 39 (1952).

[4] M. A. Gavrilov, Structural Theory of Relay Devices. [In Russian]. VZEI (1956).

[5] M. A. Gavrilov, "Determining the sequence of operation of elements in switching circuits," [In Russian]. Automation and Remote Control (USSR) 13, 5 (1952).

[6] V. N. Roginskii, "Equivalent transformations of class P switching circuits," [In Russian]. Dokl. AN SSSR 113, 2 (1957).

[7] V. N. Roginskii, "The synthesis of mixed type P switching circuits," [In Russian]. Automation and Remote Control (USSR) 18, 12 (1957).*

[8] B. I. Aranovich, "An investigation of matrix methods in questions of switching circuit structural analysis," [In Russian]. Automation and Remote Control (USSR) 10, 6 (1949).

[9] A. G. Lunts, "Algebraic methods for the analysis and synthesis of switching circuits," [In Russian]. Izv. AN SSSR, ser. matem. 16, 5 (1952).

*See English translation.

[10] M. L. Vittenberg, The Design of Telephonic and Code Relays. [In Russian]. Gosénergoizdat (1947).

[11] G. N. Povarov, "On certain matrix methods for the analysis of switching circuits," [In Russian] in Collected Works on Automation and Remote Control. [In Russian]. Izdatelstvo AN SSSR (1956).

Received May 14, 1959

EXPERIMENTAL DETERMINATION OF POWER POSITIONING SERVO SYSTEM CHARACTERISTICS

M. B. Tumarkin

(Khar'kov)

A method is considered for the investigation and testing of positioning servo systems which permits the direct recording of the systems' velocity characteristics. The distinguishing feature of this method is the application to the system's input of a forcing stimulus with constant acceleration.

Power "positioning" servo systems have, in recent years, been widely used for the automation of many types of machines. Thus, for example, several thousand machine stands are being used today in which servo systems are employed for the mechanization and automation of the processing of irregularly shaped surfaces, as well as for other purposes.

In a large proportion of the cases, such systems consist of an executive mechanism, a controlling organ and a feedback path shunting the two. In these systems the executive mechanism motion must have an aperiodic character with tractive forces of the order of hundreds of kilograms and more, with velocities of as much as several meters per minute. Under these conditions, the system must provide a high degree of accuracy in following position or translation, where the size of the discrepancies, as a function of the system's purpose, must lie approximately within the limits of 0.03 to 0.3 mm.

A very cogent question today is that of methods of testing and investigating these systems. One seeks methods which would allow the basic characterizing parameters of the systems, necessary for their rational technological use, to be determined.

Among the basic system characteristics should be placed its velocity characteristic, this being the dependence of the velocity v of the executive mechanism on the magnitude of the discrepancy h. The velocity characteristic of a system mirrors the actual features of the phenomena which occur in the system during its operation, and is capable of being used to estimate the following (tracking) error, the system's zone of insensitivity (dead zone) and the system's tractive properties (behavior of the system under loads). These data are sufficient to allow the designer and technician to judge the possibilities of using the system in concrete conditions of its employment.

From the velocity characteristic one can determine the system's velocity gain, and also the form of the system's nonlinearities, these being necessary data for the design analysis of the system.

The development of methods for taking off the velocity characteristics of power positioning servo systems is of considerable interest. Essentially, this question reduces to the choice of the character of the forcing stimulus (input function) to the system, $y(t)$, which would allow one to obtain such characteristics by the simplest and most reasonable methods.

From this point of view, let us make an estimate of several standard forms of forcing stimuli.

The unit pulse stimulus. Generation of an instantaneous unit stimulus to the controlling element is not a difficult task. For this purpose one ordinarily uses a plunger which, on priming, compresses a spring after which,

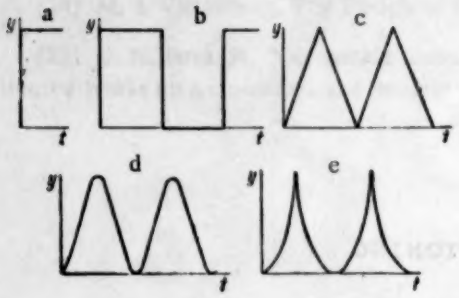

Fig. 1.

upon release of the index pin, it provides a very rapid translation of the controlling organ. It is assumed that this action is implemented instantaneously (Fig. 1a). The actual stimulus (shown by the dotted line) differs somewhat from the theoretical.

The unit pulse method is one of the standard ones for the investigation of automatic control systems [1]. It is used for taking off the characteristics of the transient response, and is capable, to a definite degree, of indicating the system's behavior from the point of view of its stability. In particular, the so-called "time constants" are determined, these entering in the form of coefficients in the differential equations describing the system.

For the investigation of the systems under consideration here, the unit pulse method is of limited applicability, and is ordinarily used for generating equal stimuli for the investigation of system stability [5]. Taking off system characteristics by this method is an extremely laborious task which, in many cases, does not achieve its aim. However, investigation of system stability by the unit pulse method is not always successful, since the system might have an oscillatory mode, not at the moment when a new position is reached, but at some intermediate position for motion of the system with an acceleration which, which such experiments, might not be expressed.

To illustrate this statement, we adduce the example of the investigation of a hydraulic servo system with a controlling valve of positive overlap. Investigation of such a system by the unit pulse method showed that it was stable, but when the system worked in the continuous tracking mode, unstable portions were observed.

Therefore, for investigating the systems under consideration here, this method is of only supplementary value.

Rectangular-pulse forcing functions. A rectangular pulse is applied periodically to the system's controlling element (Fig. 1b).

Such a method is employed in the investigation of linear (or near-linear) automatic control systems with slowly flowing processes [2].

For the systems considered here, this method is not applied; it is essentially a modification of the unit pulse method, and is of little avail in obtaining velocity characteristics of systems.

Constant-velocity forcing functions. The stimulus $y' = v = \text{const}$, carrying out its translations with uniform velocity, is applied to the controlling element. When the forcing motion is implemented by a cam, it may easily be given a periodic character (Fig. 1c). With this, the graph of the stimulus consists of segments with derivatives of constant modulus and alternating sign.

The greater part of the investigations of the servo systems under consideration here are carried out with such stimuli [3, 4, 6, 7].

In this case, one obtains the system's velocity characteristic by successively changing the forcing velocity. Each velocity, plus the corresponding discrepancy, is recorded, giving one point on the velocity characteristic.

With such a method of investigation, the accurate taking off of the system's characteristic is a very tedious and arduous experiment. When the transition is made from one velocity to another, it must be verified that all the remaining system parameters (oil pressure, temperature, load, etc) are unchanged, which is not always easily accomplished. Moreover, such an investigation gives no picture of system behavior when it moves with variable speed which, in many cases, is particularly important. The velocity characteristic cannot be directly given, but must be obtained only after the experimental data is processed; this introduces a definite subjectivity into the results obtained.

Thus, the method of investigating servo systems by forcing functions with constant velocity, although sufficiently simple, cannot be acknowledged as ideal.

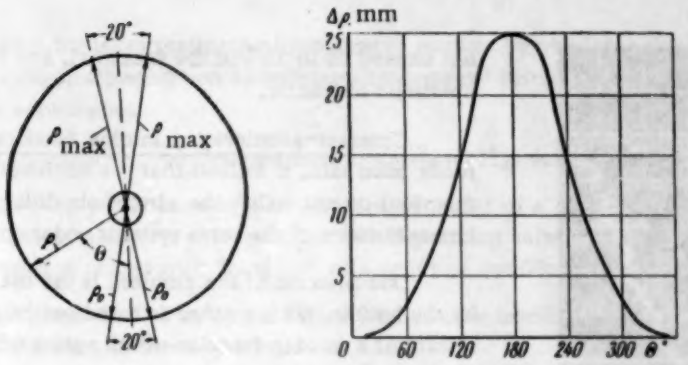

Fig. 2.

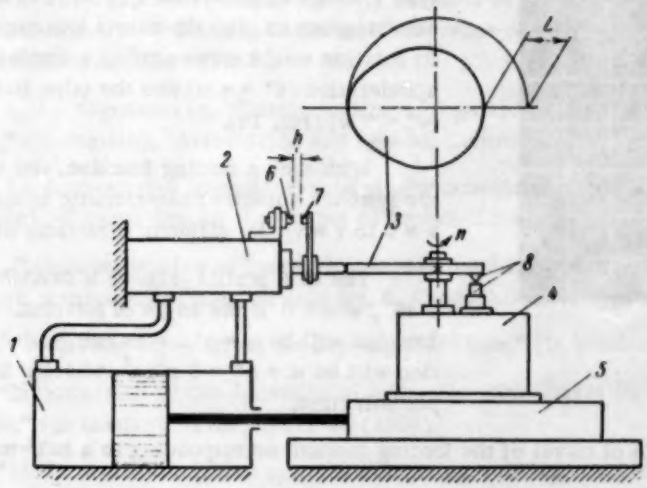

Fig. 3.

Sinusoidal forcing functions. The particularly widely used method of harmonic sinusoidal forcing functions allows one to obtain the amplitude and phase-amplitude frequency characteristics of the system, from which one can determine the system's transfer function plus a number of other characteristics, and from which one can also estimate the system's stability.

In such a method of investigation, the harmonic stimulus $y = a(1 - \cos \omega t)$ (Fig. 1d) is applied to the controlling element.

The frequency method of investigating automatic control systems is widely known and widely employed [1]. The question arises as to the effectiveness of this method for investigating servo systems of the type considered here.

It is necessary to start from the purposes for which the investigations of these systems are intended. As a result of an investigation, one must be able to obtain the system characteristic which will enable one to establish the technical possibilities of the system. The investigation must explain the influence of various factors on this characteristic, its correspondence with theoretical deductions, etc. It has already been stated that the velocity characteristic is such a characteristic, giving, as it does, the speed of tracking motion as a function of the magnitude of the discrepancy (for a given load).

With a sinusoidal input, such a characteristic can only be obtained after the proper processing of the results of an investigation. This processing requires the expenditure of a significant amount of time, and is very tedious.

In investigations of the systems considered here, the capabilities provided by this method (determination of cut-off frequency, regions of stability, etc) are not used, since the working velocity of executive motion does

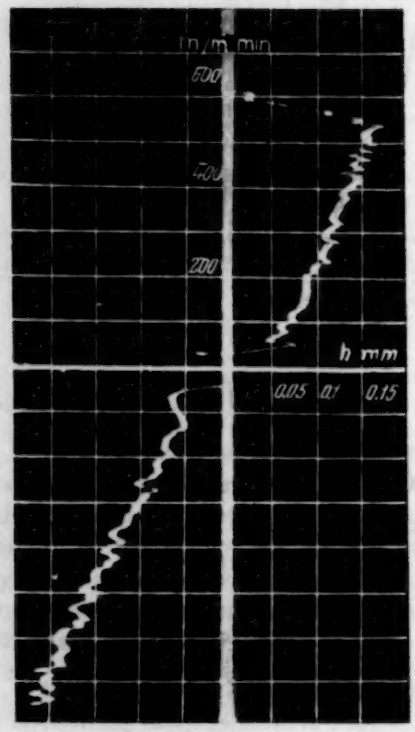

Fig. 4.

not exceed 50 to 70% of the maximal, and the motion itself has an aperiodic character.

Constant-acceleration forcing functions. From what has already been said, it follows that the enumerated types of forcing functions do not satisfy the aim of obtaining directly the velocity characteristics of the servo systems under consideration here.

One solution of the problem is the use of a method suggested by the author. This method is based on the introduction in the system of a forcing function which varies with constant acceleration, $y'' = w = \text{const}$.

The controlling element has applied to it a stimulus of the type $y = wt^2/2$, generated by a cam of the proper profile. It is advantageous to give the cam a symmetric profile, so that half of its rotation would correspond to a forcing function with positive acceleration ($y'' = +w$) and the other half, to a negative acceleration ($y'' = -w$) (Fig. 1e).

With such a forcing function, one can immediately record the system's velocity characteristic in the velocity range from $v = 0$ to $v = v_1$, for different directions of motion.

The cam profile (Fig. 2) is described by the curve $\rho = \rho_0 + a\theta^2$, where θ is the angle of rotation. The velocity of the forcing function will be $v = \rho'\omega = 4\pi an\theta$, and the forcing function acceleration will be $w = \rho'' = 8\pi^2 an^2$, where n is the number of cam rotations per unit time.

If we take the length of travel of the forcing motion, corresponding to a half-rotation of the cam, as equal to L, then

$$a = \frac{L}{\pi^2}, \qquad v = \frac{4}{\pi} Ln\theta, \qquad w = 8Ln^2.$$

By varying the number of rotations, n, of the forcing function, one varies the range of velocities of the forcing function, within which the velocity characteristic is taken off. There is, with this, a corresponding increase in the acceleration of the forcing function, which agrees with the actual conditions of system operation, since an increase in speed of the executive motion in a servo system ordinarily is accompanied by an increased acceleration. This is related to the requirement that accuracy of tracking (following) be maintained.

If it is necessary to increase the velocity of the forcing function without a proportional increase in the acceleration, one changes the cam profile by varying the angle corresponding to the necessary, or chosen, travel, L, of the forcing motion.

With a given profile and a known number of cam rotations, it is not necessary to record the forcing function. It suffices to note, on the oscillogram, the beginning of motion, when $v = 0$, and the termination, when $v = v_{max}$. In this case, the time marking on the oscillogram gives, in the proper scale, the velocity of motion. The system's velocity characteristic will be fixed on the oscillogram by recording the magnitude of the error (discrepancy). The oscillogram will also give the system's dead zone, and an indication of the character of its motion, whether aperiodic or oscillatory.

Figure 3 gives a functional schematic of the equipment for investigation of systems by the method described here.

The plane cam 3, with the profile considered above, acts on controlling element 2, which controls servomotor 1. This latter translates carriage 5, on which is positioned mechanism 4 which transmits the rotation of cam 3. The magnitude of the discrepancy h is recorded by transducer 6-7. Cam rotation is noted by marker 8. Mechanism 4 regulates the number of cam rotations for obtaining a given speed range.

The method given here for investigating positional servo systems by obtaining the speed characteristics directly is an objective method, allowing one to estimate the systems' behavior and operations in conditions which approximate to the actual ones.

The oscillogram of a system's speed characteristic is shown on Fig. 4.

The oscillogram shows the presence of a dead zone (of the order of ± 0.05 mm) in the system's characteristic, this dead zone being related to the operation of a separating valve in the system. The oscillogram also shows that system operation occurs with "hunting" as a result of insufficient system damping.

The method presented here can be used, not only for laboratory investigations, but also under industrial conditions for testing and finishing of servo systems.

LITERATURE CITED

[1] V. A. Trapeznikov and B. Ya. Kogan, "Contemporary methods of experimental investigation of automatic control systems," [In Russian]. Trudy Second All-Union Conference on Automatic Control Theory, volume III- [In Russian]. Izdatelstvo AN SSSR (1956).

[2] L. M. Krassov and A. A. Tagaevskaya, "Determination of regulator amplitude-phase characteristics by the rectilinear wave method," [In Russian]. Automation and Remote Control (USSR) 14, 3 (1953).

[3] A. F. Domrachev, "A comparative investigation of single-coordinate hydraulic servo systems used in copying machines," [In Russian]. Dissertation for the degree of "uchenoi stepeni kand. tekhn. nauk." (Kiev, 1954).

[4] B. L. Korobochkin, "An investigation of hydraulic copying machines with a constant resulting feed," [In Russian]. Raboty laboratorii stankostroitel'nogo zavoda im. S. Ordzhonikidze, vyp. 1, TsBTI (1955).

[5] V. A. Leshchenko, "Hydraulic servo drives for copying machines," [In Russian]. Trudy NIAT 92 (1957).

[6] Ya. A. Kamenir, "Investigation of the dynamics of hydraulic servo drives for copying machines for working complicated surfaces," [In Russian]. Trudy NIAT 95 (1958).

[7] M. B. Tumarkin, "Modernization of servo systems of copying-milling machines," [In Russian]. Stanki i Instrument 11 (1957).

Received August 7, 1958

THE USE OF MAGNETIC AMPLIFIERS IN INTEGRATING CIRCUITS

S. B. Negnevitskii

(Moscow)

Two basic circuits employing magnetic amplifiers as integrators are considered.

Approximate expressions are derived for the transfer functions of these amplifiers, both without the amplifier time constant being taken into account, and with it.

The disadvantages and advantages of integrating magnetic amplifiers are given.

Cases frequently arise in technology when it is necessary to integrate dc voltages. As is well known, passive RC and RL circuits can be used for this purpose. However, the capabilities of passive circuits are very limited.

In practice, therefore, active integrating devices are ordinarily employed, namely, the so-called integrating amplifiers. The most widely-used of these is the integrating amplifier with a parallel feedback path, the block schematic of which is shown in Fig. 1.

The expression for the transfer function of such an integrating amplifier is usually given with the assumptions that the input impedance R_0 and the amplifier's natural time constant τ_a are equal, respectively, to infinity and zero. When vacuum-tube amplifiers are used, such assumptions are very close to reality. If magnetic amplifiers (MAs) are used, their comparatively small input impedance and their time constants should be taken into account.

By initially taking only the input impedance into account, we obtain the following expression for the transfer function of the integrating amplifier, constructed in accordance with the circuit of Fig. 1:

$$W(p) = \frac{U_{out}}{U_{in}} = \frac{k_u}{1 + \frac{R}{R_0}} \frac{1}{1 + Tp}. \qquad (1)$$

Here, the amplifier's voltage gain is $k_u = U_{out}/U_0$, the system's time constant is

$$T = \frac{1 + k_u}{1 + \frac{R}{R_0}} \tau \qquad (2)$$

and the loop time constant is $\tau = RC$.

By substituting $R_0 \to \infty$ in (1) and (2), we obtain the well-known expression for a vacuum-tube integrating amplifier. It is clear from (1) and (2) that taking the input impedance into account does not changes the form of the transfer function, but leads only to a decrease in the system's time constant and, by this, to an increase in the error of integration (for the same time of integration).

The voltage gain is not always the characteristic feature of magnetic amplifiers. It is advantageous to introduce for them the concepts of current gain

$$k_i = \frac{U_{\text{out}}}{I_0} \tag{3}$$

and magnetizing force gain

$$k_F = \frac{U_{\text{out}}}{F_{\text{in}}} = \frac{U_{\text{out}}}{I_0 W_{\text{in}}}. \tag{4}$$

The coefficients k_u, k_I and k_F are interrelated by the expressions

$$k_u = \frac{k_i}{R_0}, \quad k_F = \frac{k_i}{W_{\text{in}}}, \quad k_u = \frac{k_F W_{\text{in}}}{R_0}. \tag{5}$$

With this, the time constant of the system, (2), can be written in the form

$$T = \frac{R_0 + k_i}{R_0 + R} \tau. \tag{6}$$

Ordinarily, in the practical applications of MAs, $k_I \gg R_0$. With this,

$$W(p) = \frac{k_i}{R + R_0} \frac{1}{1 + Tp}, \tag{7}$$

where

$$T \approx \frac{k_i C}{1 + \frac{R_0}{R}}. \tag{8}$$

For Tp $\gg$ 1, the system approximates to an ideal integrator with transfer function

$$W(p) \approx \frac{1}{RCp}. \tag{9}$$

Moreover, it is frequently the case that $R \gg R_0$, and then

$$T \approx k_i C. \tag{10}$$

It is interesting to note that the numerical value of the coefficient k_I, with the dimensions of ohms, corresponds to the input impedance of the amplifier in a circuit with 100% negative, fixed, series-connected feedback.

The special features of an MA allow one to implement an integrating MA with no galvanic connection between input and output. With this, one applies to one winding the input voltage or current, proportional to the quantity to be integrated and, by means of the second winding, one implements the negative feedback (Fig. 2). An integrating amplifier constructed in accordance with the circuit of Fig. 2 will be called an integrating amplifier with magnetic feedback; its transfer function is written in the form

$$W(p) = \frac{U_{\text{out}}}{F_{\text{in}}} = k_F \frac{1 + p\tau}{1 + pT}. \tag{11}$$

Here

$$\tau = R_0 C, \quad T = (R_0 + W_0 k_F) C. \tag{12}$$

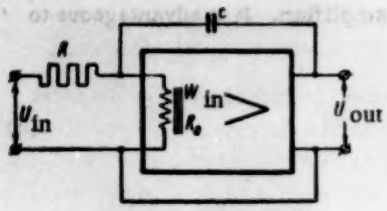

Fig. 1.

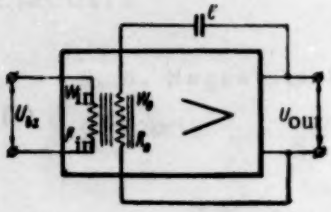

Fig. 2.

With $p\tau \ll 1$, which ordinarily is the case, the expression for the system's transfer function takes the same form as for the circuit with parallel feedback (1):

$$W(p) \approx k_F \frac{1}{1+pT}. \tag{13}$$

For $pT \gg 1$, when

$$W(p) \approx \frac{k_F}{pT}, \tag{14}$$

the system approximates to an ideal integrator.

In practice, the following relationship ordinarily holds

$$W_0 k_F = k_i \gg R_0, \tag{15}$$

and then expression (12), for the time constant T of an integrating amplifier with magnetic feedback, coincides with (10); with this, one has in expression (15) the current gain by the feedback winding W_0.

It should be mentioned that, in vacuum-tube integrating amplifiers, in order to provide negative feedback it is necessary that the polarity of the output voltage be the reverse of the polarity of the voltage at the amplifier's input. This explains why, in practice, for changing the sign of the output voltage of a vacuum-tube amplifier, an additional special inverting amplifier is connected. In an integrating MA with magnetic feedback, the output voltage may have any polarity necessary, since a negative feedback can be provided by the proper connection of the terminals of the feedback winding.

This property is an advantage of integrating MAs with magnetic feedback, particularly for use in complex interconnected circuits.

We now consider the approximate effect of the amplifier's natural time constant τ_a. We assume that the amplifier consists of two links, an amplifying one and an inertial one, and has the transfer function

$$W_a(p) = \frac{k_u}{1+p\tau_a}.$$

The transfer function of the integrating amplifier constructed in accordance with the circuit of Fig. 1, with τ_a taken into account, will have a comparatively complex expression; however, in practice, the relationships $\tau \gg \tau_a$ and $k_u/(1 + R/R_0) \gg 1$ usually hold, and then, approximately

$$W(p) = \frac{k_u}{1+\frac{R}{R_0}} \frac{1}{(1+Tp)\left(1+\frac{\tau_a}{k_u}p\right)}. \tag{16}$$

As should be expected, the presence of τ_a leads to the appearance of a second-order term in the denominator of the transfer function.

One easily obtains from (16) the zone of possible integrating frequencies. For $\tau_a = 0$, this zone lies between the limits of $1/T$ and ∞, and for $\tau_a \neq 0$, the upper frequency will be bounded by the quantity k_u/τ_a. It is important to emphasize that, as is clear from (16), the effect of the time constant τ_a is to decrease proportionately the gain. This circumstance widens the possibilities of MA employment in integrating circuits.

It should not be assumed that an integrating MA with magnetic feedback is best in all cases. Depending on the concrete conditions; internal impedance of the transducer, the possibility of connecting the transducer to a matched load, the limits of variation of the input voltage, the amplifier parameters, etc., it is advantageous to use one integrating circuit or another.

The detailed make-up of both integrating MA circuits is not the problem of the present paper. As an example, we consider a case where the problem is only solved by the use of a MA with magnetic feedback.

Let there be a transducer with maximal emf of $E_{in} = 0.02$ volt and internal impedance of $R_i = 10$ ohms. We assume that on the magnetic amplifier, with account taken of the least possible diameter of the winding conductor, one input winding can be mounted, with $W_{in} = 6350$, $R_0 = 950$ ohms, $k_1 = 952500$ ohms, or two input windings can be mounted, with $W_{in\,1} = 300$, $R_1 = 10$ ohms, $k_{11} = 45000$ ohms and $W_{in\,2} = 5000$, $R_2 = 750$ ohms and $k_{12} = 750000$ ohms.

The amplifier magnetizing force gain is $k_F = 150$.

The input-output characteristic of the MA is linear up to the limit of $U_{out} = 70$ volts.

It is required to obtain a system with time constant $T = 200$ seconds and maximal output voltage of $U_{out} = 45$ volts during an integration time of $t = 200$ seconds.

We consider the possibilities of using an amplifier with parallel, and with magnetic, feedback.

Integrating Amplifier with Parallel Feedback

As the input winding, we use the winding with $W_{in} = 6350$ and $R_0 = 950$ ohms, and we assume that the loop impedance R is also equal to 950 ohms. Since, in the given case, the condition that $k_1 \gg R_0$ holds we can, by using (8), determine the necessary magnitude of capacitance: $C = 420$ microfarads.

With a step-function connection of the maximum input signal, the output voltage after an integration time of 200 seconds will, in correspondence with (9), equal

$$U_{out} \approx E_{in} \frac{t}{RC} \approx 10 \text{ v}$$

If a lower value of R is taken, a larger value of capacitance C will be required, while if a larger magnitude for R is chosen, the already small value of output voltage is decreased.

Thus, the given MA, used as an integrating amplifier in accordance with the circuit of Fig. 1, can provide the required magnitude of T with a comparatively large capacitance C, but cannot, in the given conditions, provide the necessary U_{out}.

Integrating Amplifier with Magnetic Feedback

In this case, the amplifier must have two input windings. As the input winding for the transducer we take the winding with $W_1 = 300$ and, as the feedback winding, $W_2 = 5000$. Since relationship (15) holds, the necessary value of capacitance can be found from (10): $C = 266$ microfarads.

We now determine the input magnetizing force:

$$F_{in} = \frac{E_{in}}{R_i + R_1} W_1 = 0.3aw.$$

With a step-function connection of the input voltage, the output voltage after the integration time will, on the basis of (14), equal

$$U_{out} = E_{in} \frac{t}{W_0 C} = 45\text{v}.$$

It is clear from this example that, with the given concrete conditions and for one and the same input quantities, the use of the integrating amplifier with magnetic feedback allows one to use a capacitance smaller by a factor of 1.6, and to obtain an output voltage greater by a factor of 4.5 than in the case of parallel feedback.

Experimental work with integrating MAs shows that the assumptions made are valid, and gives a comparatively good agreement between experimental data and results computed from the formulae in this paper.

SUMMARY

Magnetic amplifiers have a host of specific special features, thanks to which integrating MAs have their disadvantages and advantages in comparison with vacuum-tube amplifiers. One of the fundamental disadvantages of integrating MAs is the decrease in the band of possible integrating frequencies, since an MA has a comparatively large time constant. Moreover, in an integrating amplifier with parallel feedback, the system time constant is decreased by a factor of $(1 + R/R_0)$; this decrease will be significant in cases where an MA is used, since its input impedance R_0 is comparatively small.

The advantages of an integrating MA are the greater reliability of operation, the lower expenditure of energy, the lower influence of the condenser bleeder resistance in the feedback circuit and of the impedance of the isolation circuit, since the comparatively low input impedance of the MA causes a correspondingly low value of impedance in the remaining circuits. In an integrating MA with magnetic feedback, the polarity of the output voltage is not rigidly related to the polarity of the input voltage. In such an amplifier, the input quantity may be applied from a current generator, while an integrating amplifier with parallel feedback is useless in principle in such a case. The input impedance, in comparison with all other schemes, has a comparatively low value and can be easily matched with the transducer's internal impedance. By means of an integrating amplifier with parallel feedback, one can integrate the sum of several signals from different sources without interconnecting them galvanically. The absence of any galvanic connection between amplifier input and output simplifies its use in complicated interconnected circuits.

Received December 7, 1958

CHRONICLE

SCIENTIFIC SEMINAR ON PNEUMOHYDRAULIC AUTOMATA

A. I. Semikova

The Third All-Union Conference-Seminar on Pneumohydraulic Automata was held May 18-19, 1959 at the Institute for Automation and Remote Control of the Academy of Science of the SSSR.

The conference was attended by 308 people from 18 cities; a total of 98 organizations working in the domain of pneumo-automata were represented.

In opening the session, seminar director M A. Aizerman noted the important role played by the annual conferences in matters of organization of work and in becoming cognizant of new developments in the given domains of science. Very important also are the personal acquaintanceships made during the course of these seminars.

The conference was opened by a plenary session at which several papers were presented, after which the work of the seminar was carried in two sections: pneumo-automata and hydro-automata; a concluding plenary session was open to general discussion.

The first paper presented was that of Z. Ya. Beirakh on an electronic-pneumatic autoregulation system in which electrical measuring and summing devices were joined with a pneumatic executive mechanism which is more reliable than an electromechanical one. As transducers and amplifiers, the electronic portion of the system control, by means of an electropneumatic relay, a piston pneumatic servomotor.

The system permits the automation of production where there is a risk of fire or explosion, since all electrical elements and circuits can be removed to a significant distance from the object itself and from the controlling organs.

V. D. Mironov (Moscow), in the paper, "Sampled-data electronic-hydraulic regulators," spoke of the electronic-hydraulic variants of mass-produced sampled-data electronic-mechanical control systems. The speaker provided descriptions of two variants of sampled-data electrohydraulic executive mechanisms. The first variant contains a unique power element, namely, a geared hydroturbine controlled by a flexible valve. The second variant (simplified) is a membrane executive mechanism, controlled by two electrovalves.

The paper of A. A. Tale and T. K. Berends (Moscow), "New elements of digital pneumatic devices," contained descriptions of developed system elements which permit the construction of complex pneumatic switching circuits and any finite automata. The pneumovalves and pneumorelays permit the implementation of the elementary operations of instantaneous logic (negation, repetition, conjunction, etc) and any primitive (non-sequential) circuit. Moreover, the one-cycle delay elements constructed from this same apparatus show the possibility, in conjunction with the instantaneous elements, of being used in the construction of nonprimitive (sequential) circuits. The continuous memory cell which has been developed has an independent value in its own right.

Based on the elements developed, devices were constructed which have important industrial value: isodrome regulators with automatic changeover, controlling color-detectors (chromographs), etc.

E. S. Arzumanov (Kirovokan) spoke of multiple-projection pneumo-electrical and pneumatic recording devices.

The Section on Pneumo-Automata

The first speaker in the section of pneumo-automata was L. A. Zalmanzon (Moscow) who, with the paper, "On the characteristics of small-size intervalve chambers," discussed an investigation of the characteristics of chambers in which, due to the constriction of the air flow by the walls, the pressure along the chamber length is significantly changed. The paper contained a generalization of the characteristics, and cited a method for designing certain types of chambers. There was mention of the possibilities of using small intervalve chambers as correctors or as transformers of one pressure as a function of another.

L. V. Vaiser (Moscow), in the paper, "Analysis of the possibilities of conversion of pneumo-automata instruments to lower supply pressures," considered the possibility of converting pneumo-automata instruments to supply pressures of 100 mm of water, which would provide a number of advantages, both technological and economical.

The paper provided the results of experimental testing and comparison of the operation of basic pneumosystem elements with supply pressures of 1 atmosphere and of 100 mm of water. Considered were "nozzle-flapper" type relays, membrane blocks, dead-end and flow-through chambers and long lines.

The results of the experimental investigation allowed the conclusions to be drawn that conversion to supply pressures of 100 mm of water would not impair the operation of pneumosystem elements and, in addition, would essentially simplify the instruments' air supply.

P. M. Shanturin (Moscow) provided information on the design development of pneumo-automata instruments at the design branch of the TsLA, particularly of the design development of pneumatic instrument elements made of plastic compositions.

Plastic composition pieces, prepared by pressing, have a heightened sturdiness and are readily amenable to mechanical processing. Mass-production assembly of parts is greatly simplified, since testing and fitting are reduced. Instruments prepared of plastic compositions can be used in chemically active (corrosive) media, are frost-resistant and heat-resistant and, moreover, their preparation is economically advantageous.

V. G. Ponomarenko (Khar'kov), in the paper, "The use of pneumatic control apparatus for automating the benzene-scrubbing branch of a coal-tar chemical factory," described a scheme for automating the benzene section and the operation of a connected-control device which, with quality transducers present, allows a quality-control system to be constructed and which permits the unifying of control of all parameters in one aggregate.

In the communication on their joint work by N. L. Kharas, V. R. Anders and T. K. Berends (Moscow), there was a description of a pneumatic control block of a controlling chromograph designed for the continuous inspection of the gas flux make-up directly at the technological stands. An industrial chromograph, supplied with a pneumatic control block, opens new possibilities for complex automation. The control block was constructed of pneumatic digital apparatus. It contains a timer (a command instrument) and an executive device (two disjunction circuits and a continuous memory cell).

A. A. Tagaevskaya (Moscow), in her communication, spoke of a developed attachment to an extremal regulator for its use on objects in which the quantity to be optimized depends on the position of n controlling organs. To realize the attachment, one used elements from systems of pneumatic discrete logical devices.

A. L. Semikova (Moscow), in the communication, "Experimental investigation of the characteristics of jet elements of pneumo-automata," revealed the results of an experimental investigation of the effect of changing elements' geometric dimensions on their output characteristics.

R. A. Auzan (Moscow) gave a paper on the subject, "Dyamic characteristics of pneumatic regulators and their adjustment on industrial objects." The paper presented the results of an experimental investigation of the dynamic characteristics of mass-produced pneumatic control apparatus.

The data of this work allowed one to characterize and judge the dynamic quality of standard blocks – the unified system aggregate (USA), the small unified system aggregate (SUSA), the system of the Chermet Central Laboratory for Automation (CLA) – and to develop instructions for the adjustment of USA regulators for a large class of industrial objects. The results were illustrated on one of the standard USA blocks. Estimates were given of the speed of response of several old type transducers and compensated transducers.

N. D. Lanin (Moscow), in his paper, spoke of pneumatic analog computers. The operational elements used in the machine do not have any limitations on integrating time, for all practical purposes.

Circuits were given for a differentiating link and for a delay block.

V. I. Chernyshev spoke of work, carried out jointly with V. N. Dmitriev, on the investigation of a piston pneumatic executive mechanism with a constant load on the piston rod. The communication contained the presentation of a graphicoanalytic method for computing the speed of a piston pneumatic executive mechanism with a constant load on the piston rod. The speaker provided a comparison of experimental and computed data.

V. M. Gorokhov (Khar'kov) gave a paper on the subject, "A piston pneumatic servo drive for general industrial use." The new design of the pneumatic drive differs from the well-known ones by its simplicity, compactness and high sensitivity.

In a joint communication, V. V. Volkov, E. Yu. Gutnikov and M. P. Kogon (Sverdlovsk) spoke of the automation of long-travel pneumatic drives.

In the communication of G. T. Berezovets on his joint work with post-graduate student Chzhou Tszin-Len (Moscow), the results were given of an experimental study of the characteristics of membrane pneumatic regulators. The experimentally-determined characteristics were given of the translation of rigid center and of the change in pressure drop as functions of changes in effective membrane area, as well as the results of a study of membrane operation for membranes prepared of different materials. An estimate was given of the effect of membrane characteristics on the quality of pneumatic device operation .

A. I. Kazeev (Kirovokan), in his communication, gave recommendations on methods of designing and choosing controlling valves.

E. A. Andreeva (Moscow) reported on the results of an investigation of power and discharge characteristics of "nozzle-flapper" type elements for the flow of compressible viscous liquids. A comparison of experimental and calculated curves was given in the communication.

In a joint communication, A. G. Arkad'ev and M. L. Podgoetskii (Moscow) spoke of pneumatic transformers, transforming small forces of 0.2 to 10 grams to standard compressed-air pressures of 0.2 to 1 atmosphere, and of the development of a pneumatic transducer with an isodrome amplifier in the feedback line. The transducer has a dynamic tuning element.

The Section of Hydro-Automata

V. N. Veller (Moscow), in the paper, "Theory and design of high-sensitivity piston mechanisms," presented a solution to the problem of increasing the sensitivity of piston mechanisms by the use of self-centering pistons. By means of special flow rings and drainage flanges on the piston's working surface, the piston is maintained strictly coaxial with the cylinder during operation.

L. Yu. Klugman (Odessa), in the paper, "Certain considerations in the choice of hydro-copying machine systems," considered hydro-copying machines and their basic elements, namely, valves, executive mechanisms and energy sources (pumps).

The quality of a system is defined by the ratio of the cylinder diameters of the given system to the diameter of the cylinder of the system of minimal dimensions. Recommendations on the choice of hydro-copying systems were given.

Yu. E. Zakharov, in the communication, "On hydrodynamic force and the discharge coefficient in valves," spoke of computing the hydrodynamic force and the discharge coefficient by the method of "jet flow" theory.

A. F. Arkhangel'skii spoke of a developed small-size pump with high productivity of high pressures for the hydrosystem of a 250 hp tractor.

In his address, M. B. Tumarkin presented a uniform approach to the study of hydraulic systems by three steps:

1) system consideration is carried out from the points is view both of its design and of the control of the executive mechanisms;

2) the system is judged with respect to six basic characteristics (speed, traction, speed gain, error, coefficient of rigidity and coefficient of insensitivity);

3) system testing is carried out by taking down the characteristic in the form of the complete curve when a signal with constant acceleration is applied.

In his address, N. S. Gamynin gave the characteristics of hydraulic aggregates used in servo systems. The speaker provided the velocity, frequency and load characteristics, as well as the transfer functions, of hydrosystem elements (valves, pump drives) and of the systems as wholes.

The paper of V. M. Dvoretskii (Moscow) was devoted to a device for obtaining small stable discharges of fluids, the principle of operation being control by the pressure drop on a constant valve.

V. D. Konstantinov, in the communication, "A hydro-aggregate as the object of synchronous generator rotational speed control," considered a hydro-aggregate consisting of a drive and a synchronous generator, gave various designs, deduced the differential equations and estimated the lag time.

In the concluding plenary session, G. I. Papernii (Leningrad) presented a paper, "A pneumatic system for the automatic control of combustion and supply processes in boilers," in which the speaker dealt with a pneumatic system, developed in the TsKTI, for the automatic control of the combustion and supply processes in boilers. The development of the pneumosystem was based on the principle of force compensation, as being the most promising.

Industrial testing under actual conditions of use of the pneumatic automatic control system demonstrated the efficiency of all its elements, as well as their high sensitivity.

The paper of V. L. Korobochkin (Moscow) contained the results of an investigation into the dynamics of hydraulic drives of copying machines.

In conclusion, there was an address by V. S. Prusenko (Moscow) with some examples of the construction of scheme of cascaded and interconnected control of thermal processes based on the instruments of a pneumatic unified system aggregate designed for general industrial use.

The comrades who took part in the discussion were unanimous in noting the great value of holding annual conferences, a number of remarks and desires were expressed for better organization of the exchange of experience among the organizations taking part in the work of the conference.

Several comrades, namely, P. M. Shanturin (Moscow), A. F. Arkhangel'skii (Chelyabinsk) and L. A. Sopochkin (Lisichansk), expressed their desires that the organization of the annual conferences be improved, that the publication of the Proceedings (Trudy) of the conference be accelerated, that there be a preliminary distribution of the subjects of the papers, that exhibits be organized, that there be a special bulletin of our seminar.

Participants I. F. Kozlov (Moscow), V. G. Ponomarenko (Khar'kov) and others also expressed the desire for a broader enlistment in the work of the annual conferences of workers in factories and related branches of industry, on whom depends the realization of the designs developed.

The chief designer of the "Tizpribor" factory, M. I. Zhutovskii, noted the parallelism in the work of different organizations, and expressed his opinion on the necessity for the coordination of work. The work of coordination should be headed up, in his opinion, by the IAT AN SSSR and TsNIIKA.

A. P. Kopelovich stated the opinion that it is necessary to coordinate, with the consumer, the technical conditions on the instruments developed up to the time they are mass-produced.

Z. Ya. Beirakhom also expressed the desire to bring KB* closer to the factories, or to strengthen the KB at the factories by changing the system of payment.

B. F. Stupak mentioned the necessity of organizing the issuance of catalogs by factories-producers.

It was proposed, moreover, that the last two pages of the journal "Priborostroenie" be reserved for a different type of advertisement and informational material.

L. L. Feigel'son, G. I. Papernyi, Yu. E. Zakharov and others then spoke of their work and noted the inadequacies of the currently existing information on current work and results obtained. In connection with this, the great value of these annual conferences was once again asserted.

In conclusion, the conference turned to letters to the TsK KPSS on questions related to the present state and prospective development of pneumohydraulic automata.

*Publ. Note: KB, Design Office.